

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:41:17 ; Search time 23 seconds
(without alignments)
2558.852 Million cell updates/sec

Title: US-10-045-072-2

Perfect score: 5788

Sequence: 1 MSTHTSSTLPAPFKILVANR.....RVVPAATKVEGDLIVVVS 1140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/aa/5A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/aa/5B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5788	100.0	1140	3	US-09-220-081-2
2	5788	100.0	1140	4	US-09-677-575-2
3	3483	60.2	1124	4	US-08-311-731A-10
4	2464	42.6	1154	4	US-09-134-001C-3428
5	2463.5	42.6	1163	4	US-09-134-000C-5707
6	2174	37.6	973	4	US-09-107-532A-4810
7	1192	20.6	694	4	US-09-433-043B-126
8	1063	18.7	456	4	US-09-634-238-276
9	1043.5	18.0	593	4	US-09-433-043B-122
10	1036.5	17.9	447	1	US-08-611-107-6
11	1036.5	17.9	447	2	US-08-422-560A-6
12	1036.5	17.9	447	3	US-08-468-793-6
13	1004.5	17.4	453	1	US-08-611-107-8
14	1004.5	17.4	453	2	US-08-422-560A-8
15	1004.5	17.4	453	3	US-08-468-793-8
16	1002.5	17.3	453	4	US-09-433-043B-121
17	995.5	17.2	453	1	US-07-956-700B-6
18	995.5	17.2	453	1	US-08-476-537-6
19	995.5	17.2	453	1	US-08-485-607-6
20	995.5	17.2	453	2	US-08-475-879-6
21	995.5	17.2	453	4	US-09-433-043B-6
22	967.5	16.7	474	4	US-09-328-352-7562
23	966.5	16.7	1116	4	US-09-252-991A-24374
24	961	16.6	453	4	US-09-543-681A-5871
25	952	16.4	605	4	US-09-433-043B-123
26	944.5	16.3	451	4	US-09-540-236-3047
27	938	16.2	448	1	US-08-074-121-3

28	938	16.2	448	5	PCT-US94-06447-3	Sequence 3, Appli
29	937	16.2	427	1	US-07-956-700B-3	Sequence 3, Appli
30	937	16.2	427	1	US-08-476-537-3	Sequence 3, Appli
31	937	16.2	427	1	US-08-485-607-3	Sequence 3, Appli
32	937	16.2	427	2	US-08-475-879-3	Sequence 3, Appli
33	937	16.2	427	4	US-09-433-043B-3	Sequence 3, Appli
34	935.5	16.2	454	4	US-09-198-452A-197	Sequence 197, App
35	935.5	16.2	701	4	US-09-252-991A-27999	Sequence 27999, A
36	926	16.0	536	3	US-08-662-344-2	Sequence 2, Appli
37	924.5	16.0	485	4	US-09-252-991A-25980	Sequence 25980, A
38	924	16.0	449	1	US-08-074-121-6	Sequence 6, Appli
39	924	16.0	449	5	PCT-US94-06447-6	Sequence 6, Appli
40	914	15.8	457	4	US-09-134-000C-4554	Sequence 4554, Ap
41	909	15.6	676	4	US-09-252-991A-26143	Sequence 26143, A
42	899	15.5	670	4	US-09-328-352-6725	Sequence 6725, Ap
43	896	15.5	619	4	US-09-543-681A-5852	Sequence 5852, Ap
44	891	15.4	652	4	US-09-328-352-5587	Sequence 5587, Ap
45	889	15.4	573	4	US-09-328-352-6420	Sequence 6420, Ap

ALIGNMENTS

RESULT 1

US-09-220-081-2
; Sequence 2, Application US/09220081
; Patent No. 6171833
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/220,081
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-220-081-2

Query Match	100.0%	Score 5788	DB 3	Length 1140
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1140	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	MSTHTSSTLPAPFKILVANRGEIAVRAFAALETGAATVAIYPRDRGSGFHRSEAVR	60	
Db	1	MSTHTSSTLPAPFKILVANRGEIAVRAFAALETGAATVAIYPRDRGSGFHRSEAVR	60	
Qy	61	IGTEGSPVKAYLDIDIIIGAAKKVKADAIYPGYPFLSENAQLARECAENGITFIGTPEV	120	
Db	61	IGTEGSPVKAYLDIDIIIGAAKKVKADAIYPGYPFLSENAQLARECAENGITFIGTPEV	120	
Qy	121	LDITGDKSRAVTAAKKAGLPVLAESTPSKNIDIVKSAEGQYPIFVKAVAGGGGGMRF	180	
Db	121	LDITGDKSRAVTAAKKAGLPVLAESTPSKNIDIVKSAEGQYPIFVKAVAGGGGGMRF	180	
Qy	181	VASPDRLKLTATASREAAAFGDDGAVYVERAVINPQHIEVQILGHTGEVHLIERDCS	240	
Db	181	VASPDRLKLTATASREAAAFGDDGAVYVERAVINPQHIEVQILGHTGEVHLIERDCS	240	
Qy	241	LQRHOKVVEIAPQALHDLPELRDRIKADAVKFCRSIGYOGAGTVEFLVDEKGNHVIEMN	300	
Db	241	LQRHOKVVEIAPQALHDLPELRDRIKADAVKFCRSIGYOGAGTVEFLVDEKGNHVIEMN	300	
Qy	301	PRIQEHTVTTEVTEVDLVKAQVRLAAGATLKELGLTQDKIKTHGAALQCRITTEPNNG	360	
Db	301	PRIQEHTVTTEVTEVDLVKAQVRLAAGATLKELGLTQDKIKTHGAALQCRITTEPNNG	360	
Qy	361	FRPDTCTITAYRSPGGAGVLDGAAQLGGELIAHFPSMLVKMTCRGSDFETAVARAQAL	420	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:44:22 ; Search time 50 Seconds
(without alignments)
5904.182 Million cell updates/sec

Title: US-10-045-072-2
Perfect score: 5788
Sequence: 1 MSTHTSTLPAPKKILVANR.....RVVPAATKVEGDLIVVWS 1140

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US02_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5788	100.0	1140	9	US-09-974-973-19
2	5788	100.0	1140	9	US-09-738-626-4265
3	5788	100.0	1140	14	US-10-045-072-2
4	5789	99.5	1157	9	US-09-974-973-2
5	5789	99.5	1157	9	US-09-974-973-4
6	4647.5	80.3	1141	12	US-10-282-122A-53885
7	3791.5	65.5	1127	12	US-10-282-122A-61950
8	3712.5	64.1	1127	12	US-10-282-122A-62684
9	3712.5	64.1	1127	12	US-10-282-122A-64810
10	2552	44.1	1144	12	US-10-282-122A-52777
11	2526	43.6	1148	12	US-10-282-122A-46866
12	2525.5	43.6	1150	15	US-10-369-493-17338
13	2524	43.6	1148	12	US-10-369-493-23112
14	2493	43.1	1146	12	US-10-282-122A-60488
15	2491.5	43.0	1144	12	US-10-282-122A-52088

16	2491.5	43.0	1144	15	US-10-369-493-7766
17	2488	43.0	1147	9	US-09-815-242-5468
18	2488	43.0	1150	12	US-10-282-122A-44391
19	2472.5	42.7	1142	9	US-09-815-242-10806
20	2471.5	42.7	1147	15	US-10-369-493-11450
21	2471.5	42.7	1151	15	US-10-369-493-14817
22	2471.5	42.7	1151	15	US-10-369-493-15006
23	2470.5	42.7	1152	15	US-10-369-493-12027
24	2464.5	42.6	1142	12	US-10-282-122A-57942
25	2464.5	42.6	1175	15	US-10-369-493-6504
26	2463.5	42.6	1142	12	US-10-282-122A-42528
27	2457	42.4	1178	15	US-10-369-493-21939
28	2443	42.2	1167	15	US-10-369-493-22819
29	2443	42.2	1185	15	US-10-369-493-2488
30	2436	42.1	1180	15	US-10-369-493-1491
31	2406	41.6	1137	15	US-10-369-493-18351
32	2375	41.0	1143	12	US-10-282-122A-52917
33	2299	39.7	1073	9	US-09-815-242-12361
34	2185.5	37.8	1385	15	US-10-369-493-3805
35	2096.5	36.2	903	15	US-10-369-493-14300
36	1827.5	31.6	1076	15	US-10-369-493-13831
37	1377.5	23.8	1122	12	US-10-282-122A-46406
38	1297	22.4	641	12	US-10-282-122A-57483
39	1290.5	22.3	611	12	US-10-282-122A-70702
40	1285.5	22.2	611	12	US-10-282-122A-70190
41	1276.5	22.1	611	12	US-10-282-122A-71785
42	1245.5	21.5	596	15	US-10-369-493-8430
43	1212.5	20.9	716	15	US-10-369-493-12942
44	1210.5	20.9	581	15	US-10-369-493-19178
45	1156	20.0	454	15	US-10-369-493-13206

ALIGNMENTS

RESULT 1

US-09-974-973-19
; Sequence 19, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteri
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; PRIOR FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 19
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-19

Query Match	100.0%	Score 5788	DB 9	Length 1140
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1140	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MSTHTSTLPAPKKILVANRGEIAVAFRAALLETGAATVAIYPREDRGSFHRSFASEAVR	60	
DB	1	MSTHTSTLPAPKKILVANRGEIAVAFRAALLETGAATVAIYPREDRGSFHRSFASEAVR	60	
QY	61	IGTEGSPVKAYLDIDETIGAAKVKADAIYPGKGFSENAQLARECAENGITFIGPTPEV	120	
DB	61	IGTEGSPVKAYLDIDETIGAAKVKADAIYPGKGFSENAQLARECAENGITFIGPTPEV	120	
QY	121	LDLTGDSRAVTAAKKAGLPVLAESTPSKNIDIVKSAEGQTYPIFKAVAGGGGRMRF	180	
DB	121	LDLTGDSRAVTAAKKAGLPVLAESTPSKNIDIVKSAEGQTYPIFKAVAGGGGRMRF	180	
QY	181	VASPELRKLATASRAEAFGDAVYVZRAVINPOHIEVQILGDHTGWHLYERDCS	240	

```

181 VASPDRLKLAATESREAAAFGDCGVVVERAVINPQHIEVQILGDHTGBVHLYERDCS 240
241 LQRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGVQAGTVEFLVDEKGNHVFIEKN 300
241 LQRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGVQAGTVEFLVDEKGNHVFIEKN 300
301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNN 360
301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNN 360
361 PRDPTGTTITAYRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQAL 420
361 PRDPTGTTITAYRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQAL 420
421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480
421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480
481 TVNKPBGVPRKDVAAPIDKLPNIKOLPLPRGSRDLKQLGPAAFARDLRQDALAVTDTT 540
481 TVNKPBGVPRKDVAAPIDKLPNIKOLPLPRGSRDLKQLGPAAFARDLRQDALAVTDTT 540
541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAMGGATYDVAMRFLFEDPDWRLDE 600
541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAMGGATYDVAMRFLFEDPDWRLDE 600
601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCFKFAVKEAASSGVDFRIFDALNDVSQMRPA 660
601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCFKFAVKEAASSGVDFRIFDALNDVSQMRPA 660
661 IDAVLETTNTAVAEVAMVAYSGDLSNENKLYTLDVYKMAEIEIVKSGAHILAKDMAGLLR 720
661 IDAVLETTNTAVAEVAMVAYSGDLSNENKLYTLDVYKMAEIEIVKSGAHILAKDMAGLLR 720
721 PAAVTKLVTLRREFDLPHVHTHTAGGQATLTAFAAQAQADAVDGNASPLSGTTTQPS 780
721 PAAVTKLVTLRREFDLPHVHTHTAGGQATLTAFAAQAQADAVDGNASPLSGTTTQPS 780
781 LSAIVAAFAHTRRDGLSLAEVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
781 LSAIVAAFAHTRRDGLSLAEVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
841 SNLRAQATLGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVGDLALHLVAGVDPAF 900
841 SNLRAQATLGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVGDLALHLVAGVDPAF 900
901 AADPOKYDIPDSVIAFLRGELGNPPGWPPELRTALEGRSEGGKAPLTFEVEEQAHLDA 960
901 AADPOKYDIPDSVIAFLRGELGNPPGWPPELRTALEGRSEGGKAPLTFEVEEQAHLDA 960
961 DSKERRNSLRLFPKPTTEFLHRRRFGNTSALDDREFFYGLVEGRETILRLPDVTRTP 1020
961 DSKERRNSLRLFPKPTTEFLHRRRFGNTSALDDREFFYGLVEGRETILRLPDVTRTP 1020
1021 LLVRLDAISEPDDKGMNVVANNVQIRPMRVDRSVESVTATAEKDSSNKGHVAAPFA 1080
1021 LLVRLDAISEPDDKGMNVVANNVQIRPMRVDRSVESVTATAEKDSSNKGHVAAPFA 1080
1081 GVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140
1081 GVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140

```

RESULT 2
 US-09-738-626-4265
 ; Sequence 4265, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO

```

; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4265
; LENGTH: 1140
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4265

Query Match 100.0%; Score 5788; DB 9; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHTHSSSTLPARKKILVANRGEIATRAALETGAATVAIYPRDRGSPHRSFASAVR 60
DB 1 MSHTHSSSTLPARKKILVANRGEIATRAALETGAATVAIYPRDRGSPHRSFASAVR 60
QY 61 ICTEGSPVKAYLIDIEIIIGAAKKVKADAIYPGYGFSLSENAQLARECAENGITFTTPEV 120
DB 61 ICTEGSPVKAYLIDIEIIIGAAKKVKADAIYPGYGFSLSENAQLARECAENGITFTTPEV 120
QY 121 LDLTGDSKAVTAAKKAGLPVLAESTPSKNIDEIVKSEAGQTYPIFVKAVAGGGGRWRF 180
DB 121 LDLTGDSKAVTAAKKAGLPVLAESTPSKNIDEIVKSEAGQTYPIFVKAVAGGGGRWRF 180
QY 181 VASPDRLKLAATESREAAAFGDCGVVVERAVINPQHIEVQILGDHTGBVHLYERDCS 240
DB 181 VASPDRLKLAATESREAAAFGDCGVVVERAVINPQHIEVQILGDHTGBVHLYERDCS 240
QY 241 LQRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGVQAGTVEFLVDEKGNHVFIEKN 300
DB 241 LQRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGVQAGTVEFLVDEKGNHVFIEKN 300
QY 301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNN 360
DB 301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNN 360
QY 361 PRDPTGTTITAYRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQAL 420
DB 361 PRDPTGTTITAYRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQAL 420
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480
DB 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480
QY 481 TVNKPBGVPRKDVAAPIDKLPNIKOLPLPRGSRDLKQLGPAAFARDLRQDALAVTDTT 540
DB 481 TVNKPBGVPRKDVAAPIDKLPNIKOLPLPRGSRDLKQLGPAAFARDLRQDALAVTDTT 540
QY 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAMGGATYDVAMRFLFEDPDWRLDE 600
DB 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAMGGATYDVAMRFLFEDPDWRLDE 600
QY 601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCFKFAVKEAASSGVDFRIFDALNDVSQMRPA 660
DB 601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCFKFAVKEAASSGVDFRIFDALNDVSQMRPA 660
QY 661 IDAVLETTNTAVAEVAMVAYSGDLSNENKLYTLDVYKMAEIEIVKSGAHILAKDMAGLLR 720
DB 661 IDAVLETTNTAVAEVAMVAYSGDLSNENKLYTLDVYKMAEIEIVKSGAHILAKDMAGLLR 720

```

Db 661 IDAVLENTAAVAEMAYSGDLSFNEKLYTLDYLLKVAEEIVKSGAHILAIDKMAAGLLR 720
Qy 721 PAAVTKLVATLREEDLPVHVHTHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSPS 780
Db 721 PAAVTKLVATLREEDLPVHVHTHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSPS 780
Qy 781 LSAIVAAFAHTRDRDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQOL 840
Db 781 LSAIVAAFAHTRDRDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQOL 840
Qy 841 SNLRQAATLGLADRELEEDNVAANVEMGRPTKVTSSKVVGDALHLVAGVDPADF 900
Db 841 SNLRQAATLGLADRELEEDNVAANVEMGRPTKVTSSKVVGDALHLVAGVDPADF 900
Qy 901 AADPKYDIPDSVIAFLRGELGNPPGWPPEPLRTRALEGRSEKAPLTVPEEQAHLDA 960
Db 901 AADPKYDIPDSVIAFLRGELGNPPGWPPEPLRTRALEGRSEKAPLTVPEEQAHLDA 960
Qy 961 DSKERRNSLNRLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db 961 DSKERRNSLNRLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Qy 1021 LLVRLDAISEPDDKGMENNVANVNGQIRPMVRDRSVESVTATAEKAADSSNKGHVAAAPPA 1080
Db 1021 LLVRLDAISEPDDKGMENNVANVNGQIRPMVRDRSVESVTATAEKAADSSNKGHVAAAPPA 1080
Qy 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140

RESULT 3

US-10-045-072-2
; Sequence 2, Application US/10045072
; Publication No. US20030027305A1
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; TITLE OF INVENTION: Pyruvate Carboxylase from *Corynebacterium glutamicum*
; FILE REFERENCE: 1533.0730002
; CURRENT APPLICATION NUMBER: US/10/045,072
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 09/677,575
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: *Corynebacterium glutamicum*
US-10-045-072-2

Query Match 100.0%; Score 5788; DB 14; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSHTSSTLPAFKKILVANKEGEIAVRAALETGAATVAIYPREDRSGSFHRSFASAVR 60
Db 1 MSHTSSTLPAFKKILVANKEGEIAVRAALETGAATVAIYPREDRSGSFHRSFASAVR 60
Qy 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGPGFLSENAQARECAENGITFPGTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGPGFLSENAQARECAENGITFPGTPEV 120
Qy 121 LDLTGDKSRVATAKAGLPVLAESTPSKNIDEVKSAEGQTYPIFKVAVAGGGGGMRF 180
Db 121 LDLTGDKSRVATAKAGLPVLAESTPSKNIDEVKSAEGQTYPIFKVAVAGGGGGMRF 180
Qy 181 VASPELRLKULTEASREAEAFDGAAGVYVERAVINPOHIEVQILGHTGEVHLVERDCS 240
Db 181 VASPELRLKULTEASREAEAFDGAAGVYVERAVINPOHIEVQILGHTGEVHLVERDCS 240

RESULT 4

US-09-974-973-2
; Sequence 2, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from *Corynebacteri*
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21

Db 181 VASPELRLKULTEASREAEAFDGAAGVYVERAVINPOHIEVQILGHTGEVHLVERDCS 240
Qy 241 LORRHQKWEIAPQAHLDPBLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVEIEMN 300
Db 241 LORRHQKWEIAPQAHLDPBLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVEIEMN 300
Qy 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGHTQDKIKTHGAALOCRIITTEPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGHTQDKIKTHGAALOCRIITTEPNNG 360
Qy 361 FRPDTGTITAYRPGGAGVRLDGAALGGEBITAHFDSMLVMKTCRGSDFETAVARAQAL 420
Db 361 FRPDTGTITAYRPGGAGVRLDGAALGGEBITAHFDSMLVMKTCRGSDFETAVARAQAL 420
Qy 421 AEFTVSGVANIGFLRALUREEDFTSKRIATGFIADHPHLLQAPPADDEGCRILDLVADV 480
Db 421 AEFTVSGVANIGFLRALUREEDFTSKRIATGFIADHPHLLQAPPADDEGCRILDLVADV 480
Qy 481 TVNPKHGVRRPKDVAAPIDKLPNIKDLPLPRGSDRLKQLGPAAFARDLRQDALAVTDTT 540
Db 481 TVNPKHGVRRPKDVAAPIDKLPNIKDLPLPRGSDRLKQLGPAAFARDLRQDALAVTDTT 540
Qy 541 FRDAHQSILLATRVRSFALKPAEAEAVAKLTPELLSVEAWGATYDVAMRFLPEDPWDLDE 600
Db 541 FRDAHQSILLATRVRSFALKPAEAEAVAKLTPELLSVEAWGATYDVAMRFLPEDPWDLDE 600
Qy 601 LREAMPNNVQMLLRGNTVGYTPYDPSVCRAFAVKEAASGVDFRIFDALNDVQMRPA 660
Db 601 LREAMPNNVQMLLRGNTVGYTPYDPSVCRAFAVKEAASGVDFRIFDALNDVQMRPA 660
Qy 661 IDAVLENTAAVAEMAYSGDLSFNEKLYTLDYLLKVAEEIVKSGAHILAIDKMAAGLLR 720
Db 661 IDAVLENTAAVAEMAYSGDLSFNEKLYTLDYLLKVAEEIVKSGAHILAIDKMAAGLLR 720
Qy 721 PAAVTKLVATLREEDLPVHVHTHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSPS 780
Db 721 PAAVTKLVATLREEDLPVHVHTHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSPS 780
Qy 781 LSAIVAAFAHTRDRDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQOL 840
Db 781 LSAIVAAFAHTRDRDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQOL 840
Qy 841 SNLRQAATLGLADRELEEDNVAANVEMGRPTKVTSSKVVGDALHLVAGVDPADF 900
Db 841 SNLRQAATLGLADRELEEDNVAANVEMGRPTKVTSSKVVGDALHLVAGVDPADF 900
Qy 901 AADPKYDIPDSVIAFLRGELGNPPGWPPEPLRTRALEGRSEKAPLTVPEEQAHLDA 960
Db 901 AADPKYDIPDSVIAFLRGELGNPPGWPPEPLRTRALEGRSEKAPLTVPEEQAHLDA 960
Qy 961 DSKERRNSLNRLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db 961 DSKERRNSLNRLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Qy 1021 LLVRLDAISEPDDKGMENNVANVNGQIRPMVRDRSVESVTATAEKAADSSNKGHVAAAPPA 1080
Db 1021 LLVRLDAISEPDDKGMENNVANVNGQIRPMVRDRSVESVTATAEKAADSSNKGHVAAAPPA 1080
Qy 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140

Thu Mar 25 09:51:47 2004

;; PRIOR APPLICATION NUMBER: US 60/239,913
;; PRIOR FILING DATE: 2000-10-13
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 2
;; LENGTH: 1157
;; TYPE: PRT
;; ORGANISM: Corynebacterium glutamicum
US-09-974-973-2

Query Match 99.5%; Score 5759; DB 9; Length 1157;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPAPFKKILVANRGEIAVRAFAALLETGAATVAIYPREDRGSFHRSPASEAVR 60
DB 18 VSTHTSSTLPAPFKKILVANRGEIAVRAFAALLETGAATVAIYPREDRGSFHRSPASEAVR 77

QY 61 IGTGSPVKAYLDIDEIIGAARKKADAIYPCYGFLENAQLARECAENGITFIPTPEV 120
DB 78 IGTGSPVKAYLDIDEIIGAARKKADAIYPCYGFLENAQLARECAENGITFIPTPEV 137

QY 121 LDLTGDKSRVTAATAKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGGGGRMRF 180
DB 138 LDLTGDKSRVTAATAKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGGGGRMRF 197

QY 181 VASPDRLKRLATEASAEAAFGDGNVYVERAVINPQHIEVQILGDHTGEVHLVERDCS 240
DB 198 VSSPDLRLKRLATEASAEAAFGDGSVYVERAVINPQHIEVQILGDHTGEVHLVERDCS 257

QY 241 LQRRHCKVVEIAPAOHLDELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300
DB 258 LQRRHCKVVEIAPAOHLDELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 317

QY 301 PRIQVHTVTEVTEVDLVKAQMRLAAGATLDELGTQDKIKTHGAALOCRIITDPNNG 360
DB 318 PRIQVHTVTEVTEVDLVKAQMRLAAGATLDELGTQDKIKTHGAALOCRIITDPNNG 377

QY 361 FRPDTGTTITAYRSPGGAGVRLDGAALGGIEITAHFDSMLVMTCRGSDFTETAVARAQAL 420
DB 378 FRPDTGTTITAYRSPGGAGVRLDGAALGGIEITAHFDSMLVMTCRGSDFTETAVARAQAL 437

QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPADDEQRIILDYADV 480
DB 438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPADDEQRIILDYADV 497

QY 481 TVNKPVGVRPKDVAAPIDKLPNIDPLPRGSRDLKQLPAAFPARDLREQDALAVDTT 540
DB 498 TVNKPVGVRPKDVAAPIDKLPNIDPLPRGSRDLKQLPAAFPARDLREQDALAVDTT 557

QY 541 FRDAHQSLLATRVRSFALKPAEAVAKLTPELLSVEANGGATYDVAMEFLFEDPDWRLDE 600
DB 558 FRDAHQSLLATRVRSFALKPAEAVAKLTPELLSVEANGGATYDVAMEFLFEDPDWRLDE 617

QY 601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFVKEAASSGVDFIRIPDALNDVSYQMRPA 660
DB 618 LREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFVKEAASSGVDFIRIPDALNDVSYQMRPA 677

QY 661 IDAVLETNATAEAVAMAYSGDLSDPNEKLYTLDYLYKVAEETVKSAGHILAIKOWAGLLR 720
DB 678 IDAVLETNATAEAVAMAYSGDLSDPNEKLYTLDYLYKVAEETVKSAGHILAIKOWAGLLR 737

QY 721 PRAVTKLVTLAREFPLPVHVTHTDTAGGQLATYFAAAQAGADAVDGSAPLSGTTSPQS 780
DB 738 PRAVTKLVTLAREFPLPVHVTHTDTAGGQLATYFAAAQAGADAVDGSAPLSGTTSPQS 797

QY 781 LSAIVAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVYRHEIPGQOL 840
DB 798 LSAIVAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVYRHEIPGQOL 857

QY 841 SNLRAQATLGLADREFLEDNDYAAVNEMLGRPTKVTPSSKVGVGDIALHLVGAGVDPADF 900
DB 858 SNLRAQATLGLADREFLEDNDYAAVNEMLGRPTKVTPSSKVGVGDIALHLVGAGVDPADF 917

QY 901 AADPQKYDIPDSVIAFLRGEIAGNPPGWPBPLRTRALEGRSEKAPLTEVPEEEQAHLD 960
DB 918 AADPQKYDIPDSVIAFLRGEIAGNPPGWPBPLRTRALEGRSEKAPLTEVPEEEQAHLD 977

QY 961 DSKERRNSLNRLFFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
DB 978 DSKERRNSLNRLFFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037

QY 1021 LLVRLDAISEPDDKMRNVVANVNGQIRPMKVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
DB 1038 LLVRLDAISEPDDKMRNVVANVNGQIRPMKVRDRSVESVTATAEKADSSNKGHVAAPFA 1097

QY 1081 GVVTVTAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
DB 1098 GVVTVTAEGDEVKAGDAVAIIEMKMEATITASVDGKIERVVVVPAATKVEGGDLIVVVS 1157

RESULT 5
US-09-974-973-4
; Sequence 4; Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanks, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteriu
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-4

Query Match 99.5%; Score 5759; DB 9; Length 1157;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPAPFKKILVANRGEIAVRAFAALLETGAATVAIYPREDRGSFHRSPASEAVR 60
DB 18 VSTHTSSTLPAPFKKILVANRGEIAVRAFAALLETGAATVAIYPREDRGSFHRSPASEAVR 77

QY 61 IGTGSPVKAYLDIDEIIGAARKKADAIYPCYGFLENAQLARECAENGITFIPTPEV 120
DB 78 IGTGSPVKAYLDIDEIIGAARKKADAIYPCYGFLENAQLARECAENGITFIPTPEV 137

QY 121 LDLTGDKSRVTAATAKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGGGGRMRF 180
DB 138 LDLTGDKSRVTAATAKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGGGGRMRF 197

QY 181 VASPDRLKRLATEASAEAAFGDGNVYVERAVINPQHIEVQILGDHTGEVHLVERDCS 240
DB 198 VSSPDLRLKRLATEASAEAAFGDGSVYVERAVINPQHIEVQILGDHTGEVHLVERDCS 257

QY 241 LQRRHCKVVEIAPAOHLDELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300
DB 258 LQRRHCKVVEIAPAOHLDELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 317

QY 301 PRIQVHTVTEVTEVDLVKAQMRLAAGATLDELGTQDKIKTHGAALOCRIITDPNNG 360
DB 318 PRIQVHTVTEVTEVDLVKAQMRLAAGATLDELGTQDKIKTHGAALOCRIITDPNNG 377

QY 361 FRPDTGTTITAYRSPGGAGVRLDGAALGGIEITAHFDSMLVMTCRGSDFTETAVARAQAL 420
DB 378 FRPDTGTTITAYRSPGGAGVRLDGAALGGIEITAHFDSMLVMTCRGSDFTETAVARAQAL 437

QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPADDEQRIILDYADV 480
DB 438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPADDEQRIILDYADV 497

```
QY 481 TVNPKGVKPVAAPIIDKLPNIKOLPLPRGRDRDLKOLGPAAFARDLREQDALAVDTDTT 540
Db 498 TVNPKGVKPVAAPIIDKLPNIKOLPLPRGRDRDLKOLGPAAFARDLREQDALAVDTDTT 557
QY 541 FDAHOSLLATVRSPALKPAEAAVAKLTPELLSVEAWGGATYDVAMFLFEDPWRDLDE 600
Db 558 FDAHOSLLATVRSPALKPAEAAVAKLTPELLSVEAWGGATYDVAMFLFEDPWRDLDE 617
QY 601 LBREAMPNNIOMLLGRNTVGTTPYDPSVCRAFVKEAASSGVVDIIFRIDALNDVSQMRPA 660
Db 618 LBREAMPNNIOMLLGRNTVGTTPYDPSVCRAFVKEAASSGVVDIIFRIDALNDVSQMRPA 677
QY 661 IDAVLETTNTAAEAVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHLAIKDMAGLLR 720
Db 678 IDAVLETTNTAAEAVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHLAIKDMAGLLR 737
QY 721 PAAVTKLVTALRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGSAPLSGTTSPQS 780
Db 738 PAAVTKLVTALRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGSAPLSGTTSPQS 797
QY 781 LSAIVAAPFAHTRDGLSLEAVSDLEPYWEAVRGYLYLPESGTPGPTGRVYRHEITPGQL 840
Db 798 LSAIVAAPFAHTRDGLSLEAVSDLEPYWEAVRGYLYLPESGTPGPTGRVYRHEITPGQL 857
QY 841 SNLRAQATAGLADREFELIEDNYAAVNMELGRPTKVTSSKVVGDALHLVAGVDPADP 900
Db 858 SNLRAQATAGLADREFELIEDNYAAVNMELGRPTKVTSSKVVGDALHLVAGVDPADP 917
QY 901 AADPKQYDIPDSVIAELRGELNPGGWPPEPLRTALEGRSGEKAPLTVPEEBOAHLDA 960
Db 918 AADPKQYDIPDSVIAELRGELNPGGWPPEPLRTALEGRSGEKAPLTVPEEBOAHLDA 977
QY 961 DSKKERRNSLNRLLFPKPTFEELEHRRRFGNTSALDDREFFYGLVGEGRTEILRLPDVTRP 1020
Db 978 DSKKERRNSLNRLLFPKPTFEELEHRRRFGNTSALDDREFFYGLVGEGRTEILRLPDVTRP 1037
QY 1021 LLVRLDAISEPDDKMGNVVANNVNGQIRPMRVDRDSVESVTATAEKAASSNKGHVAAFPFA 1080
Db 1038 LLVRLDAISEPDDKMGNVVANNVNGQIRPMRVDRDSVESVTATAEKAASSNKGHVAAFPFA 1097
QY 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKVEATITASVDGKIIDRVVVPAAATKVEGGDILVVVS 1140
Db 1098 GVVTVTVAEGDEVKAGDAVAIIEAMKVEATITASVDGKIIDRVVVPAAATKVEGGDILVVVS 1157

RESULT 6
US-10-282-122A-53885
; Sequence 53885: Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITPA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

```
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 53885
; LENGTH: 1141
; TYPE: PR1
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-53885

Query Match      80.3%; Score 4647.5; DB 12; Length 1141;
Best Local Similarity 80.8%; Pred. NO. 0;
Matches 917; Conservative 81; Mismatches 137; Indels 3; Gaps 2;

QY 2 STHTSSTLPFAFKILVANRGEIAVRAAFRAALETGAATVAIYPREDRSGSFHRSFASAVRI 61
Db 6 SERGSTVNVLSKILVANRGEIAVRAAFRAAFETGAATVAIYPREDRSGSFHRSFASAVLI 65
QY 62 GTGSPVKAVLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITTFIGPTPEVL 121
Db 66 GEGSAVKAYLDIDEIRAAKQTDGADAIYPGYGFLSENAQLARECAENGITTFIGPTPEVL 125
QY 122 DLTGDKSRVTAAKKAGLPVLAESTTSKNIDEIVKSAEGOTYPIFKVAVAGGGRGRFV 181
Db 126 ELTGDKXAAVTAAREAGLPTLTETEATDDPKLAELTKGQTYPLFKVAVAGGGRGRFV 185
QY 182 ASPDELKLAATESRBAEAAFGGAVYVERAVINPOHIEVQILGDHTGEVHHYERDCSL 241
Db 186 DAPENLEKLAESAAREAAAAGDGRVYERAVINPOHIEVQILGDHAGNIHLHYERDCSL 245
QY 242 QRRHKQVVEIAPAHQHDPELRDRI CADAVKFCRSIGVQAGTVEFLVDEKGNHVFEMNP 301
Db 246 QRRHKQVVEIAPAHQHLAPGLREKICADAVAFARIHQVQAGTVEFLVDEKGNHVFEMNP 305
QY 302 RIQVEHTVTEBEVTLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGF 361
Db 306 RIQVEHTVTEBEVTQVLDLVKSQIMIASGATLEDGLRQEDIHTEGAALQCRITTEDPNNGF 365
QY 362 RPDGTGITAVRSPGAGVRLDGAOLGGEITAHFDSMLVMTCRGSDFFETAVARAQALA 421
Db 366 RPDGTGITAVRSPGAGVRLDGAOLGGEITAHFDSMLVMTCRGADFFATAVARAQBALA 425
QY 422 EFTVSGVATNIGFLRALLREDEFTSKRIATGCTFIADHPHLLQAPPADDEQGRILDYADVT 481
Db 426 EFTVSGVATNIGFLRALLREDEFTSKRIATGCTFIADHPHLLQAPPADDEQGRILNYADVT 485
QY 482 VNKPHGVKPVKDVAAPIIDKLPNIKOLPLPRGRDRDLKOLGPAAFARDLREQDALAVDTDTT 541
Db 486 VNKPHGLRPA-VINPVEKLPABIKGELPRGRDRDLKOLGPEGFARLARKQDALAVDTDTT 544
QY 542 RDAHQSLLATVRSPALKPAEAAVAKLTPELLSVEAWGGATYDVAMFLFEDPWRDLDEL 601
Db 545 RDAHQSLLATVRSPALKPAEAAVAKLTPELLSVEAWGGATYDVAMFLFEDPWRDLDEL 604
QY 602 REAMPNNIOMLLGRNTVGTTPYDPSVCRAFVKEAASSGVVDIIFRIDALNDVSQMRPAI 661
Db 605 REAMPNNIOMLLGRNTVGTTPYDPSVCRAFVKEAASSGVVDIIFRIDALNDVSQMRPAI 664
QY 662 DAVLETTNTAAEAVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHLAIKDMAGLLR 721
Db 662 DAVLETTNTAAEAVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHLAIKDMAGLLR 721
```

665 DAVLENTTIAEVAMAYSGDLTNPSEKLYTLDYLLKLABEIVKSGAHVLAIKDMAGLMKP 724
722 AAVTKLVATLRRERFDLPVHVHTDITAGGQATVFAAAQAGADAVDCASAPLSTTSQPSL 781
725 AATKLVTELRFPDLFPVHVHTDITAGGQATVFAAAQAGADAVDCASAPLSTTSQPSL 784
782 SAIVAAFAHTRDTGLSLEAVSDLEPFWAARGLYLPFESGTPGTRVYRHEIPGGQLS 841
785 SAIVAAFAHTRDTGLSLEAVSDLEPFWAARGLYLPFESGTPGTRVYRHEIPGGQLS 844
842 NLRAQATLGLADREFELIEDNYAAVNMELGRPTKVTTPSSKVVGDALHLVAGVDPADFA 901
845 NLRAQATLGLADREFELIEDNYAAVNMELGRPTKVTTPSSKVVGDALHLVAGVDPADFA 904
902 ADPOKYDIPDSVIAFLRGLGNPPGWPPELRTRALEGRSEGKAPLVEPVEBQALHDAD 961
905 ADPOKYDIPDSVIAFLRGLGNPPGWPPELRTRALEGRSEGKAPLVEPVEBQALHDAD 962
962 DSKERRNSLRLPKPTESEFLHRRFGNTSALDDEFFYGLVEGRETLIRLPDVRTPL 1021
963 DRATVPTLDRLLFPKPAEFAEHRROFGDTTKLGDRAFLYGLKEKEIVIRTDSEVPM 1022
1022 LVRLDAISEPDDKGMNVVAVNNGQIRPMRVDRSVESVTATAEKADSSNKGHVAE 1081
1023 LVRLDAISEPDDKGMNVVAVNNGQIRPMRVDRSVESVTATAEKADSSNKGHVAE 1082
1082 VVTVAEGDVKAGDAVIAEAMKMENTITASVDGKIDRVVPAATKVEGGDLIVV 1139
1083 VVTVTEBEGATKAGDPVAVIAEAMKMENTITASVDGKIDRVVPAATKVEGGDLIVV 1140

RESULT 7

US-10-282-122A-61950
Sequence 61950, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELTRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61950
LENGTH: 1127
TYPE: PRT
ORGANISM: Mycobacterium avium
US-10-282-122A-61950
Query Match 65.5%; Score 3791.5; DB 12; Length 1127;
Best Local Similarity 66.4%; Pred. No. 2.5e-230;
Matches 750; Conservative 140; Mismatches 229; Indels 11; Gaps 7;
QY 14 KILVANREIAVRAFAALETGAATVAIYPREDRSGPHRSFASAEVAVRIGTEGSPVKAYLD 73
DB 4 KVLVANREIAVRAFAALETGAATVAIYPREDRSGPHRSFASAEVAVRIGTEGSPVKAYLD 63
QY 74 IDEIIGAARKYKADAIYPGYFLSENAQLARECAENGIITFIGTPPEVLDITGDKSRVTA 133
DB 64 VDEIVGTALACGADAIYPGYFLSENPDLAAACAAGITFVGPSEAEVLELTGDKSRAIAA 123
QY 134 AKKAGLPVLAESTPSKNIDEIVKSAEGCTVPIFKAVAGGGGRGMRFFVASPDELRKLA 193
DB 124 ARAAGLPVLAESTPSSTSVQELLSSAETMTFFLVKAVAGGGGRGMRVTPDGALEAIEA 183
QY 194 ASREAAAFGCGAVYVERAVINPOHIEVQILGDHTGEVHLYERDCSLQRHKKVETIAP 253
DB 184 ASREAAAFGCGAVYVERAVINPOHIEVQILGDHTGEVHLYERDCSLQRHKKVETIAP 243
QY 254 AQHLDPELRDRICADAVKFCESIGYQAGATVFLVDEKGNHVFTEMNPRIOVEHTVTEEV 313
DB 244 APNLDPALRERICADAVAFARSIGYTCAGTVEFLDDEGNHVFTEMNPRIOVEHTVTEEV 303
QY 314 TEVDLVKQAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGPRPTGTITAYRS 373
DB 304 TDVDLVSAQLRIASQTLLEEIGLSQDSVSRGAALQCRITTEDPANGPRPTGTITAYRT 363
QY 374 PGAGVRIDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVATNIG 433
DB 364 PGAGVRIDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVATNIP 423
QY 434 FLRALIREDEFTSKRIATGFIADHPHLLQAPADDEQGRILDYLDADVTVNKPFGVRPXD 493
DB 424 FLOAVLDDPDFOAGRIITTSFTEQRFQTLTARSSADRGTKILNYLDADVTVNKPFGVRP 483
QY 494 AAPIDKLVNIDPL--PRGSRDLKQLGPAFAADLREQDALAVDTTFRDAHQSLIAT 551
DB 484 -YPHDKLPDI-DLAASPAGSKQRLTELGPFRFAAWLRESGGVGTDTTFRDAHQSLIAT 541
QY 552 RYRSFALKFAAEAVAKLTPELLSVEAWGATVDVAMRFLFEDPMDRLDELREAMPNVNIQ 611
DB 542 RYRTSGLLKVAPYIARTWPTQLLSVECWGATVDVALRFLKEDPWERLTALREAMPNICLQ 601
QY 612 MLLRGRNTVGYTPYPSVCRAPVKEAASGVDIIFRFDALNDVSQMRPAIDAVLETNTAV 671
DB 602 MLLRGRNTVGYTPYPTVTATFAEATETGIDIFRFDALANNVDSMRPAIDAVRETGSAI 661
QY 672 AEVAMAYSGDLSDPNEKLYTLDYLLKLABEIVKSGAHVLAIKDMAGLRLPAATVKTVAL 731
DB 662 AEVAMAYSGDLSDPNEKLYTLDYLLKLABEIVKSGAHVLAIKDMAGLRLPAATVKTVAL 721
QY 732 RREFDLPHVHTDITAGGQATVFAAAQAGADAVDCASAPLSTTSQPSLSTTSQPSLSTTS 791
DB 722 KSRFDLPVHVHTDITAGGQATVFAAAQAGADAVDCASAPLSTTSQPSLSTTSQPSLSTTS 781
QY 792 RRDGTGLSLEAVSDLEPFWAARGLYLPFESGTPGTRVYRHEIPGGQLSRLRAQATAG 851
DB 782 EYDTGLSLEAVSDLEPFWAARGLYLPFESGTPGTRVYRHEIPGGQLSRLRAQATAG 841
QY 852 LADRFELIEDNYAAVNMELGRPTKVTTPSSKVVGDALHLVAGVDPADFAADPOKYDIPD 911
DB 842 LGDRPEDIENAVAGADAILGHLVKVTPSSKVVGDALHLVAGVDPADFAADPOKYDIPD 901
QY 912 SVTAFRLGELGNPPGWPPELRTRALEGRSEGK--APLVEPVEBQALHDADDSKERRNS 969


```
Db 902 SVTGFRLGELGDPGPGWPEFLTKALQGRGPAKPEOPLT---AEDEAALAPGAR-ROAA 957
Qy 970 LNRLLPFPKTEEBLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAIS 1029
Db 958 LNRLLPFPKTEEBLEHREYQDTSGLSANOFFVGLRQGRDEHVEL-ERGVELLIGLEAIS 1016
Qy 1030 EPDCKGRNVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAFPAGVTVVAE 1089
Db 1017 DPDERGMRVTMCTLNGQLRPVVVRDSSIAIDVPAAEKADRPNDPHIAAPPAGVTVVAE 1076
Qy 1090 GDEVKAGDAVAITEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVV 1139
Db 1077 GRQVEAGQTIATIEAMKMEAAVTSKSGKVARIANSRTAQVEGGDLIMVI 1126
```

RESULT 8

```
US-10-282-122A-62684
; Sequence 62684, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsythe, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62684
; LENGTH: 1127
; TYPE: PRF
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62684
```

```
Query Match 64.1%; Score 3712.5; DB 12; Length 1127;
Best Local Similarity 64.4%; Pred. No. 4.5e-284;
Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;
Qy 12 FKKILVANRGEIAVRAFALETTGAATVAIYPRDRSGSHRSPASAVRIGTEGSPVKAY 71
Db 2 FSKVLVANRGEIAIRAFRAAYELGVGTVAIYPIEDRNSQHRLKADESQIGDIGHPVAY 51
```

```
Qy 72 LDIDEIIGAANKVADAIPYGYFLSENQAARECAENGITFTIGTPPEVULDLTGKSRV 131
Db 62 LSVDEIVATARRAGADAIYGYFLSENPDAAAACAAAGISFYGPSAEVLEAGNSRAI 121
Qy 132 TAAKAGLPVLAESTSKNIDEIVKSAEGOTYPIFKAVAGGGRGMRVFVASDELKLA 191
Db 122 AAAREAGLPVMSASASVDELLSVAAGMPFLFKAVAGGGRGMRVGDAAALPEAI 191
Qy 192 TEASREAAAFGAGVVEEAVINPOHIEVOIIGDHTGEVVHUYERDCSLQRHQKVEI 251
Db 182 EAASREASAFGDPVTYVLEQAVINPRHIEVOIILADNLGDVILHYERDCSVQRHQKVEL 241
Qy 252 APAOHLDELDRICADAVKFCISIGYQAGAGTVEFLVDEKGNHVFTEMNRI OVEHTVTE 311
Db 242 APAPHLDAELRYKMCVDVAFARHIGYSAGTVEFLDERGEYVFTEMNRPVQVEHTVTE 301
Qy 312 EVTEVDLVKAQMLAAGATLKEIGLTQDKIKTHGAALQCSITTEDPNNGRPPTGITAY 371
Db 302 EITVDVLVASQLRIAAGETLEQLGRQEDIAPEGAALQCSITTEDFANGRPPTGRISAL 361
Qy 372 RSPGAGVRLDGAAGLGEITAHFDSMLVKMTCRGSDFETAVARAORALAEFTVSGVATN 431
Db 362 RTAGGAGVRLDGSINLGAEISPYFDSMLVKLTCTGRDLPTAVSRARRAIAEFIRGUSTN 421
Qy 432 IGFLRALLREDEFTSKRIATGFIADHPHLLQAPPADDEOQRIILDYADVTVNKPHGVPRK 491
Db 422 IPFLQAVLDDPFRAGRVTTSTFIDERPOLLUTASADRGTKILNFLADVTVNPNPYSRPS 481
Qy 492 DVAAPIDKLNINKDPLP----PRGSRDELQOLGPAAFARDLREQDALAVTDTTFRDAHQS 547
Db 482 TI-YPDDKLP---DLDRAPAPGAGSKORLVKLGEGFARWLRESAAVGVVDTTFRDAHQS 537
Qy 548 LLATRVRSFALKFAAAEAVAKLTPELLSVEAGGATYDVAMRFLFEDPWRDLDELREAMPN 607
Db 538 LLATRVTSGLSRVAPYLARTMPQLLSVECGGATYDVALLRFLKEDSWERLATLRAAMPN 597
Qy 608 UNIQMLLRGNTVGYTPYDSCVCRAFVKEAASGVDFRIFDALINDVSQNRPAIDAVLET 667
Db 598 ICLQMLLRGNTVGYTPYIVTSFAVQEATATGIDIFRIFDALNNIESMRPAIDAVRET 657
Qy 668 NTAVAEAMAYSGDLSDPNEKLYLDVYLKMAEIEVKSGAHILAIDMAGLLPAAVTKL 727
Db 658 GSAIAEVAMCYTGDLTDPGQLYLDVYLKLAQIVDAGAHVLAIKDMAGLLPAAVTKL 717
Qy 728 VTALRREFDLPHVHTHTDRTAGGQATYFAAAQAGADAVDGASAPLSOTTSQPSLSAIVAA 787
Db 718 VSALRSRFDLPVHLHTDTPGQLASVYAAWHAGADAVDGAAPLAGTTTSPALSSIVAA 777
Qy 788 FAHTERDTGLSLEAVSLEPYWEAVRGLYLPFESGTPGPTGRVYVHEIPGQQLSNLRAQA 847
Db 778 AAHTEYDTGLSLSAVCALEPYWEALRKVYAPFESGLPGPTGRVYVHEIPGQQLSNLRAQA 837
Qy 848 TALGLADRFELEDYEDYAVNEMLGRPTKPTPSSKVVGDALHLVAGVDPADFAADPKQY 907
Db 838 IALGLGDRFEEIEEYAGADRVLGRLVKVTFTSKVVGDLALALVAGVSADEFASDPARF 897
Qy 908 DIPDSVIAFLRGEIENPGGWPRLRTRALEGRSEKAPLTEVPEEQAHLDADDKERR 967
Db 898 GIPESVLGFLRGEIENPGGWPRLRTRALEGRSEKAPLTEVPEEQAHLDADDKERR 955
Qy 968 NSLNRLFPKPTBEFLESHRRFNGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1027
Db 956 ATLNRLFPKPTBEFLESHRRFNGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1014
Qy 1028 ISEPDGKMNIVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAFPAGVTVTV 1087
Db 1015 ISEPDGKMNIVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAFPAGVTVTV 1074
Qy 1088 AEGDEVKAGDAVAITEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db 1075 CVGERVGGQTIATIEAMKMEAPITAPVAGTVVERVAVSDTAQVEGGDLIVVVS 1127
```



```

RESULT 9
US-10-282-122A-64810
; Sequence 64810, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chilsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64810
; TYPE: PRT
; LENGTH: 1127
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64810

Query Match      64.1%; Score 3712.5; DB 12; Length 1127;
Best Local Similarity 64.4%; Pred. No. 4.5e-284;
Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;

QY 12 FKKILVNRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPFASEAVRIGTESPVKAY 71
DB 2 FSKVLVNRGEIAVRAFAAELVGVTVAVPYEDRNSOHRKKADESVOIGDIGHPVHAY 61

QY 72 LIDELIIGNAKVKADAIYPGFLSENQAIAEACENGITFIGTPPEVLDITGDKSEAV 131
DB 62 LSVDEIVATARRAGADAIYPGYFLSENPDLAACAAGISFVGPSEVLELAGNKSRAI 121

QY 132 TAAKAGLPVLAESTPSKNIDIVKSAEQOTYPIFKVAVAGGGGGMFVFA5PDELKLA 191
DB 122 AAAREAGLPVLAESSAPSVDELLSVAAGMPFPLFKVAVAGGGGGMFVFA5PDELKLA 181

QY 192 TEASRAEAFGGGAVYVERAVINPOHIEVQILGDHTEVHVHLYERDCSLORRQKVEI 251
DB 182 EAASRAEAFGGGAVYVERAVINPOHIEVQILGDHTEVHVHLYERDCSLORRQKVEI 241

QY 252 APAQHLDPRLRICAADAVKFCRSIGYQAGTVFELVDEKGNHVFEMNPRIQVEHTVTE 311

```

```

DB 242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLDLDERGEYVFIEMNPRVQVEHTVTE 301
QY 312 EVTEVDLYKQMRLAAGATLKEGLTQDKIKTHGAALOCRTTDPNNGFPDGTGTITAY 371
DB 302 EITDVLVASQLRIAAAGETLQGLRQEDIAHPGAALOCRTTDPNNGFPDGTGTISAL 361
QY 372 RSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDPETAVARAQALAFVTSVGVATN 431
DB 362 RTAGGAGVRLDGSTNLGAEISFPFDSMLVKLTICRGRDLPVAVSARRAIAEFRIQVSTN 421
QY 432 IGLRALLREEDFTSKRIATGFIADPHLLQAPADDEQGRILLYLADVTYKNGHGVPRK 491
DB 422 IPFLQAVLDDPDFRAGRTVTSFIDERPQLLTARASADRGTKILNPLADVTYNNPYGSPS 481
QY 492 DVAAPIDKLENIKOLPL---PGSRDLKOLGPAAPAFARDLREQDALAVDTITFEDAQS 547
DB 482 TI-YFDDKLP---DLDLRAAPGAGSKORLVKLPFGPFARWLRESAAVGVDTITFEDAQS 537
QY 548 LLATRVRSFALKAABAAVAKUTPELLSVANGGATYDVAMRFLPDPDRDLDELREAMPN 607
DB 538 LLATRVTSGLSRVAPYLARTMPQLLSVECGGATYDVALRFLKEDPWERLATLRAAMPN 597
QY 608 VNIOMLLGRNTVGYTPYVDSVCRAFKVEAASSGVVDIFRIFDALNDVSOMPAIDAVLET 667
DB 598 ICLQMLLRGRNTVGYTPYVDSVCRAFKVEAASSGVVDIFRIFDALNDVSOMPAIDAVLET 657
QY 668 NTAVAEVAMAYSGDLSDPNEKLYLDYLLKMAEEIVKSGAHILAIKDWAGLLRPAATVKL 727
DB 658 GSAIAEVAMCYTGDULTDPEQLYLDYLLKMAEEIVKSGAHILAIKDWAGLLRPAATVKL 717
QY 728 VTALAREEDLPVHVHTHTAGGLATYFAAQAQADAVDGAAPLSGTTSPSLSAIVAA 787
DB 718 VSALRSRFDLPVHLHTHTDPGGLASVYVAWHAGADAVDGAAPLSGTTSPSLSAIVAA 777
QY 788 FAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGTGTVYRHEIPGGQLSNLRAQA 847
DB 778 AAHTEYDTCGLSLSAVCALEPYWEALRVYAPESGLPGTGRVYRHEIPGGQLSNLRAQA 837
QY 848 TALGLADRELIEDNYAANEMIGRPTKVTSPKVVGLALHLVGVADPADFAADPOKY 907
DB 838 IALGLDREFEIEEAYAGADRVGLRVKVTPTSKVVGDLALALVGVADPADFAADPOKY 897
QY 908 DIPDSVIAFLRGELGNPPGCGWPEPLRTRALEGRSEKAPLTVPEEEOAHLDDADSKERR 967
DB 898 GIPESVLGFLRGELGDPGCGWPEPLRTRALAGGAAR-PTAQLAADDEALSSVGAK-RQ 955
QY 968 NSLNRLLPKPTTEEFLEHRRRGNTSALDDREFFVLVEGRETILRLPDVRFLLVRLDA 1027
DB 956 ATLNRLLPKPTTEEFLEHRRRGNTSALDDREFFVLVEGRETILRLPDVRFLLVRLDA 1014
QY 1028 ISBPDCKMNVNVANVNGQIRPMVRDRSVESVTATAKADSNGKHVAAPFAGVVTVT 1087
DB 1015 ISEPDCKMNVNVANVNGQIRPMVRDRSVESVTATAKADSNGKHVAAPFAGVVTVT 1074
QY 1088 AEGDEVKAGDAVAIEMKOMEATITASVDGKIDRVVPAATVEGGDLIVVYS 1140
DB 1075 CVGERVAGGAGTATIAMKOMEAPITAPVAGTVERVAVSDTAQVEGGDLIVVYS 1127

RESULT 10
US-10-282-122A-52777
; Sequence 52777, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chilsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

```

```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-03-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52777
; LENGTH: 1144
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-10-282-122A-52777

Query Match 44.1%; Score 2552; DB 12; Length 1144;
Best Local Similarity 45.7%; Pred. No. 2.8e-192;
Matches 526; Conservative 215; Mismatches 372; Indels 38; Gaps 14;

QY 12 FKILVANRGEIAVRAAETGAATVAIYPRDRGSHFRSFASEAVRIGTEGSPVKAY 71
DB 4 FKSVLVANRGEIAVRAAETGAATVAIYPRDRGSHFRSFASEAVRIGTEGSPVKAY 63
QY 72 LDIDEIIGAKKVKADAIYGVGFLSENACLAECAGCITFIGTPEVLDITGKSRV 131
DB 64 LNEEIIQALKGVDAIHPGVGFLAENSEFAKCKREAGIEFIGPTAEMWKLGDKIKSK 123
QY 132 TAAKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIPVKAVAGGGGRGMFVSPDELRL 190
DB 124 IVAEAGVPTIPGVQKPIKSEKALEBFARYCGYPIMLKAAAGGGGRGMFVIRTEBELISS 183
QY 191 ATEASBEAEAFDGGAVYVERAVINPOHLEVOITLGDHTGEVHLYERDCSLORRHKVVE 250
DB 184 FSKAKNEAKAFGIDDIIEKYLENPKHEVLQGLGKGNIVHLYERDCSIORRHKVIE 243
QY 251 IAPAQHLDELRLRICADAVKFCRSIGYCGAGTVEFLVDKGNHVFIEKNPQIQTHTVT 310
DB 244 FTPAFALPKKREIICNDALKIAKTGVRSAGTLEFLVDTTGNHVFIEKNPQIQTHTVT 303
QY 311 EYTEVDLVKAOMRLAAGATL--KELGL-TODKIKTHGAALOCRIITEDPNNGFRDCT 367
DB 304 EMITGDIQVQSLLIAEGYKLDSEEVGKISQESIQRGYAIOCRVTTEDPSNNFADTK 363
QY 368 ITAYRSPGGAGVRLDGA-AOLGGEITAHFDSMLVQVTCRGSDFETAVARAQALAEFTVS 426
DB 364 IEDYRTGSGFIRLDGGNGFTGVSIPYDLSLVKTTNSRTFNDAIRKISIRAIKEFKID 423
QY 427 GVATNIGFLRALREEDFTSKRTATGTFIADPHLLQAPPADDEQGRILYLDVTVKPH 486
DB 424 GVKNTIGFLNLVNHQEQFRKGQCDTNFIEKNPELFDITSKTDDEVBILKEIGKVNETH 483
QY 487 GVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTTFEDAHQ 546

```

```

DB 484 GIK-KDFDVP--TIPIVDEGLSLKGTQJILDEKGPGLVSWIKTKNKLTLTDTTNRDAHQ 540
QY 547 SLIATVRFAKPAABAVAKLTPELLSVBANGCATYDVAMRFLPDPDRDLDELEZAMP 606
DB 541 SLMATRMSVDMEFKIAKQSVLGKOLFSEMMWGGATFDVAYRFLKESPTRLEERKSLIP 600
QY 607 NVNIQMLLRGNTVGYTPYDPSVCRAFVKEAASSGVVDIIFRIFDALNDVSCMRPAIDAVLE 666
DB 601 NVLFQMLIRGANAVGYKNYPDNVIRKFIQSDADSGIDVFRIFDLSNLWKGMEVATDEVLK 660
QY 667 TMTAVAEVAMAYSGDLSDPNEKYLTDLYLWAEIEVKSAGHILAIKDMAGILRPAAVTK 726
DB 661 QN-KVAETCMYCTGDIIEYRDYKSLQYYVDLAKIEKTHGAILGKDSALLKPYAAVK 719
QY 727 LVTALRRFELPVHVHTHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSPSLSAIVA 786
DB 720 LKALKNEISIPHLHHTHTTNGVATVLMAAHAGVDIVDTAFNMSGLTSQPALNSIVA 779
QY 787 APAHTRDGLSLEAVSDLEPYWEAVRGLYLPESGTGPTGRVYRHEITPGGQLSNLRAQ 846
DB 780 ALENTDRETGLDITDMQKLSDYMSAVRPVYSQFESGLKSGSABIKYKYEIPGGQYSLKFKQ 839
QY 847 ATALGLADREFELIEDNVAAVNEMLGRPTKVTPESSKVVGDALHLVAGVDPADFAADPOK 906
DB 840 VESFGLGHKEFEVKEKYNEMGLDIIKVTPESSKVVGDALFIMVKNDLTPENIYEKAEK 899
QY 907 YDIPDSVIAFLRGLNPPGPNPEIRTRALESRGSGKAPLIEVPEE----- 953
DB 900 MAPPDSAVSYFKGMGQPMGGFFPEKLQKLKGED----PITCRPGEMLPPEDEFKIREH 955
QY 954 --EQAHLDADDKSKERNLSNLLFPKPTBEFLHRRRFGNTSALDDREFFYGLVEGRETL 1011
DB 956 LKEKHLDDATEN---DIISYALYPEVDFKYLDFLKEYGDLSHGSDVFFHGLYEGTAE 1011
QY 1012 IRLPDRVITPLLVRLDAISEPDDKGMNVANVNGQIRPMVRDRS---VESVTATAEK-A 1067
DB 1012 IELQEGKT-FIVQLSEIGKVDSEGNRAVPEINGNREIRIKDKSSLMAGNI'SNSTKMA 1070
QY 1068 DSSNGKHVAAPPAG-VVTVTVAEGDVEKAGDAVAIEMKMEATITASVDGDKIDRVVVA 1126
DB 1071 DPANKKHIGSSIPGTVIKVLVNGKDEIKEGDSLIVIEAMKMETNIVASLSGVVGSLLVKE 1130
QY 1127 ATKVEGGDLIV 1137
DB 1131 GDQVKSQQLLL 1141

RESULT 11
US-10-282-122A-46866
; Sequence 46866, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

```

PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent in version 3.1
SEQ ID NO 46866
LENGTH: 1148
TYPE: PRT
ORGANISM: Bacillus anthracis
US-10-282-122A-46866

Query Match 43.6%; Score 2526; DB 12; Length 1148;
Best Local Similarity 47.2%; Pred. No. 3.2e-190;
Matches 542; Conservative 178; Mismatches 402; Indels 26; Gaps 13;

Qy 9 LPAPFKILVANRGIARAFRAALETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPV 68
Db 4 LORIQKLVANRGIARVFRACSELGLKTVIAYSKEDSGSYHRYKADESILYVGEKKPI 63

Qy 69 KAYIDDEITGAOKKADAIYGYGFLSENAQLARECAENGITFPGTPEVLDLTGDKS 128
Db 64 DAYLDIEGIIIEIAKSNHVDIHPGYGFLSENIQPAKCEBEGIFIPGKSKHDMFGDKV 123

Qy 129 RAVTAACKAGLPVLAEST-PSKNDIEIVKSAEGQTYIFVKAVAGGGGRMRVAPDEL 187
Db 124 KARTQAQLAQIPVPGSDGFDVSLSEVKEPAEKDYDPIIILKASLGGGGRMRVIRISEL 183

Qy 188 KRLATEASREAEAFGCGAVVERAVINPOHIEVQIIGDHTGEVHLYYERDCSLQRHQK 247
Db 184 RESYNRAKSEAKAAGNDEVVEKPKHIEVQIILADEEGNVVHLYYERDCSVQRHQK 243

Qy 248 VVEIAPACHLPDLDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFIEEMPRIQVEH 307
Db 244 VVEIAPSVLSDDLRLQRICERAAVKLTKNVYNLNGTVEFLVKD-DNPFYIEVPRVQVEH 302

Qy 308 TVTEEVTEVDLVKQMLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNRPD 364
Db 303 TITEMITGVDIVQSQILITADGHALHSAKMGVPGKQEEVVVHGFATQSEVTTEDPLNNFMD 362

Qy 365 TGTITAYSPGAGVRLD-GAAQIGGETAHDPSMLVMTKCRGSDRETAVARAQALAEF 423
Db 363 TGTIMAYSGGFGVRLDTGNSPGQAVTAPYDSLLVKTWTWALTFFQAAAKHERNLKEF 422

Qy 424 TVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADVYN 483
Db 423 RIRGKITNIPLENVVKHKNLSGEYDTSFIDASPELFPFKRDRGTGKMLNIGTVTVN 482

Qy 484 KPHGVPRKDVAA-RIDKLPNI-KDLPGRSFRDLKQGPAAAFARDLREODALAVDTTTF 541
Db 483 GPPGVGKKEKIFPDARIPNVHSEPIQNGTKQILDERGADGLVKVQDQKRVLLTDTTF 542

Qy 542 RDAHQSLLATRVSRFALPAAEAVAKLTPELLSVYAEWGGATYDVAMRFLFDPMDRLDEL 601
Db 543 RDAHQSLLATRIKDLHQAIEPTARMPLNLSFAEMWGGATFDVAYRFLKEDPWERLLD 602

Qy 602 REAMPNNIOMLRGNVGTPTPDSVCRFAVKEAASSGVDIETFDALNDVSQMRPAI 661
Db 603 REKPNVLFQMLLRSSNAVGNYPNDLIQKVECSAQAGIDVPRIFDLSLNWEGMRVAI 662

Qy 662 DAVLETTNTAFAEVAMAYSGDLSDFNEKLYTLDYLLKWAEBEIVKSGAHILAIDKMAGLLRP 721
Db 663 DAVRDTG-KIAEATMCTGDIHDPMRSKYDLNYYONLAKELVSGAHILGIKMAGLLKP 721

Qy 722 AAVTKLVTLARREEDLPHVHTHTAGGQATATFAAAQAGADAVDGASAPLSGTTSPSL 781
Db 722 NAAVDLVSAALKETYSIPIHLHTDTSNGILTYTKAIEAGVDIVDVAVSMAGTSPSA 781

Qy 782 SAIVAAPAHTRRDGTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQLS 841
Db 782 NTLIYALGGNERQPDVNIIDSLEKLSHYWEDVKYAPFESGMNAPHTVEVYHMEFGQYS 841

Qy 842 NIRAQATLGLADREFELIEDNYAAVNEMLGRPKTPTSSKVGDLALHLVAGVDPADPA 901
Db 842 NLOQAKVVGGLGRFDEVKVMYRVDNMFGDIVKVTPTSSKVGDMALFMVQNHLTEQDVL 901

Qy 902 ADPOKYDIPDSVIAFLRGLCNPPGWPPELRTALEGSRSEKAPLTEVPEEEQAHLDA 961
Db 902 ERGSHMDPFGSVEMFSGDLGQPYGGPKKLOEIL---KGKPLTVRPGELLEPVDFD 957

Qy 962 DSKERR-NSLNR-----LLPKPTBEFLHRRFRFGNTSALDDREFFYGLVEGRET 1010
Db 958 ALKEELFHLGREGVTMFDVVAVALYPKVFMDYKVAELYGNVSVLDTFTTFYGNRLGEEI 1017

Qy 1011 LIRLPDVTETPLVRLDAISEBDDKGMNVANVANGQIRPMRVRDRSVESVTATAEKADSS 1070
Db 1018 DVEIEQGT-LMVKLVSIGELQPDGNRVLYLEFNGQPREIVVKDESVKATVAQRVKGNRE 1076

Qy 1071 NKGHVAAFPAG-VVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATK 1129
Db 1077 NPNHISATMPTGTVIKVVVKEGDEVKGDGSMATKMETTVQAPFNGKVKKVVVNDGDA 1136

Qy 1130 VEGGDLIV 1137
Db 1137 IOTGDLII 1144

RESULT 12
US-10-369-493-17338
Sequence 17338, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-1052052B
CURRENT APPLICATION NUMBER: US/10369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 17338
LENGTH: 1150
TYPE: PRT
ORGANISM: Bacillus halodurans
US-10-369-493-17338

Query Match 43.6%; Score 2525.5; DB 15; Length 1150;
Best Local Similarity 46.8%; Pred. No. 3.5e-190;
Matches 539; Conservative 185; Mismatches 400; Indels 27; Gaps 13;

Qy 9 LPAPFKILVANRGIARAFRAALETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPV 68
Db 4 LKNIKKLVANRGEIAIRFRACTELHIRTVAIYKEDTGAYHRYKADEAYLVGEGKKPI 63

Qy 69 KAYLDIDETIGAAKKVKAADAIYGVGFLSENAQLARECAENGITFPGTPEVLDLTGDKS 128
Db 64 EAYLDIEGIIIEIAKSNHVDIHPGYGFLSENIQPAKCEBEGIFIPGKSKHDMFGDKV 123

Qy	129	RAVTAAKAGLPLVLABST--PSKNIDIVKSAEQGYPIFVKAVAGGGGGMRFVAPDDEL	187
Db	124	QAREQAIAKANLPVPGSDGPFVSSLEDVKAFADKHGYFFIIKAAKGGGGGMRIVRASENDV	183
Qy	188	RKLATASRAEAAAFGDGAVVYERAVINPOHIEVOQLIGDHTGEVHLHYERDCSLQRHQK	247
Db	184	QESYERAKSEAKAFGNDEVYVKEFIEHPKHIEVOQLADKHGNTLHLYERDCSVQRHQK	243
Qy	248	VWETIAPAQHLDPELRDRIACADAVKFCRSGYOGAGTVEFLVEKGNHVIENMPRIOVHE	307
Db	244	VWEVAPSVLSSEVBRERICQAAYQLAENVNVYVAGTVEFLVDRBGNFYFIEVNPRIQVHE	303
Qy	308	TVTEEVTEVDLVKAQWRLAAGATL--KELGL-TQDKIKTHGAALQCRITETEDPNNGFRED	364
Db	304	TITMTVTGDIVQSOLFADGEHLHGRDLGIPQBEIVCHGYAIQSERVITEDPSNGFLPD	363
Qy	365	TGHTIAYRSPGGAGVRILD-GAALGGEIINAFUSMLVRMTCRGSDETFETAVARAQRLAEF	423
Db	364	TGRINAYRSGGFGVRLDAGNGFGQAVITPYDSSLIVKVSTWALTPEGAAAKNLWREF	423
Qy	424	TVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADVTYN	483
Db	424	RIRGIKTNIAFLENNVQHRQFLSGEYNTSFIQDTPELFVFPKEDRGTKMWSFIGETIVN	483
Qy	484	KPHGVPRKDVAAIDK--LPNIK-DIPLRGSDRLKQLGPAAPARDLREQDALATVDTT	540
Db	484	GYEGLE-KTKKPVDFRPVPVKLKSBEIPDGTQKILDQHGPEGLAKWKEQKHVLUTDIT	542
Qy	541	FRDAHOSLLATRVRSFALKPAAEAVALKTELLSVEAWGGATVDVAMRFLFEDPMDRLDE	600
Db	543	FRDAHOSLLATRVKTHDKQIAEPTARLLPNLFSMEMWGGATVDVAMRFLHEDPERULLI	602
Qy	601	LRAMPNVTOMLLRGENTGYTPYDPSVCRAFKVKAASGVDIRIFDALNDVSOMRA	660
Db	603	LRKAPNVLQMLLRASNAVGYKNYPNLLIREFVDSANAGIDVRFIFSLNWWGSKLA	662
Qy	661	IDAVLENTTAAEVAMAYSGELSDPNBKLYTLDYLMKABEIVKSGAHLIAIKDMAGLLR	720
Db	663	IEAVGEAN-KIAEATICYTGILDSSRPKYDLAYYKKLAKELEAAGAHILGIKDMAGLLK	721
Qy	721	PAAVTKLVTALRREFDLPVHVHTHTAGGOLATYFAAQAGADAVDGASAPLSGTTSQPS	780
Db	722	PEAAQJVAELKOTVTIIPVHLHTHTSGNGIFTYARAIEAGVDIVDVAVSSMAGLTSQPS	781
Qy	781	LSAIAFAAHTRDRDTGSLSEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL	840
Db	782	ANSLYALADSEKOPNVNITALEQAEFWBETRKFYAGFSGMNAFHTVEYHEMPGGQY	841
Qy	841	SNLRAQATALLADRPELIEDNVAANVEMLRPTKPTPSSKVYGDIALHALHAGVDPADF	900
Db	842	SNLQOQAKAVGLGHMNEVKNMRTVNDMFVDVVKVTTPSSKVGDWALINVQNDLTTEEYV	901
Qy	901	AADPQKYDIPDSVIAFLRGEIENLPPGGWPEPLRTRALEGRSEKAPLUTEVEBEQAHLDA	960
Db	902	YENGHKLDPDSVVVEFPEGOLGQPYQFPFKKLBQBIILKGRK---PITNRPGENMSEPIQF	957
Qy	961	DDSKER-RNSINR-----LLFPKTEEFLEHRRRFGNTSALDDREFFYGLVEGRE	1009
Db	958	EAIKEELYNKLDROVQTHSDHLSYALYFKVPWEEFRFQTFQDVSVLDTPTFFVGLAPGEE	1017
Qy	1010	TLIRLPVRTPLLYRLDAISEPDKGMNVNVANVQIIRPMVRDRDSVESVTATAEKADS	1069
Db	1018	IEVEIEQKKT-LIVKFISSLKPDQDGNRIYVYFELNGQPREVLINKDSQVKTSIISRPKADK	1076
Qy	1070	SNKGHVAAPFAG-VYTVTVAGDVBKAGDAVAIIETAMKMEATITASVDGKIDRWVPAAT	1128
Db	1077	SNFNHIGASMPGTIVKALVEKGVKQGDHLMITEMAKMETTVQAPDGEVVALHVKQGD	1136
Qy	1129	KVEGGLDIVVV	1139
Db	1137	AIOGTGDLIIIEV	1147

RESULT 13
 US-10-369-493-23112
 ; Sequence 23112, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 23112
 ; LENGTH: 1148
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 US-10-369-493-23112

Query Match.	43.6%	Score 2524,	DB 15,	Length 1148;	
Best Local Similarity	46.0%;	Pred. No. 4.6e-190;			
Matches	533;	Conservative 186;	Mismatches 392;	Indels 48;	Gaps 13;
QY	11	AFKKILVANGETI	AVRAFRAL	ETGAATVAI	YPREDRGSPHRFSASEAVRIGTEGSPVKA 70
DB	5	SIQKILVANGETA	IRIFRACTEL	NIETVAVYSKEDSGSYHYRYKAD	EAVILVGEKKPIDA 64
QY	71	YLIDETIIGA	KKVKADAI	VPGVGFI	SENQALARECAENGITITGTPPEVLDLTGDKSRA 130
DB	65	YLDIEGID	IAKKNVD	AIHPGVGFI	SENTHFARRCEEIGVIFPGPXEHLDMFGDKVKA 124
QY	131	VTAAKG	AGLPLAEST	-PSKNI	DEIVKASQGYPIPVKAVAGGGGGMFVASFDELRL 188
DB	125	REQAK	KAGIPV	PGSDGPAET	LEAVEFGQANGVPIIKASLGGGGRMAIVRSESEVKE 184
QY	190	LATEAS	EEAABAF	DGNAVY	VERAVINPQHIEVQLGDHTGEVHYLYERDCSLQRPHQKV 249
DB	185	AYERAK	EEAKA	AFNGDEN	VYVEKLIENPKHTEVQVIGDKQGNVHLPFEDCSVQRHOKVI 244
QY	250	EIAPAQ	HLDEL	RDRICAD	AVKFCRSIGYCGAGTVEPLVDEKGNHVIEMNPRIQVEHTV 309
DB	245	EVAP	SVLSPEL	RDOICEA	VALAKNVYINAGTVEFLV-ANNEFFYIEYNPRQVEHTI 303
QY	310	TEEVTE	VDLVKAQ	NR	LAAGATL--KEILGLTQDK- IKTHGAALQCRITTEDPNNGFRPDTG 366
DB	304	TEMITG	VDIVQ	TQLVLAQ	GHSLASKVNIPEQKDI FTIGVIAOSRVVTEPQNDFMPEDTG 363
QY	367	TITAYR	SPGGAGV	RLD-GA	ALGGEITAHEDSLMVKMTCRDSDPETA VAAQALAEFTV 425
DB	364	KIMAYR	SGGGFGV	RLDTG	TGNSFQGANITPYDLSLVKLSTWALTQEAAAKMVRNLQEFRI 423
QY	426	SGVATN	-GFLR	ALLRESDFT	SKRIATGFIAHPHLLQAPPADDEQGRILDYLVADTVN-- 483
DB	424	RGIKNT	IPFL	ENAVAKHEK	FTLGTQYDTSFIDTTPLELFPFKQKRGTKMLTYIGVNTVNGF 483
QY	484	-----	-KPHGVR	PKDVAAP	IDKLPNIKDLPPLRGSSDRLLKOLGPAAFAFARDLRQDALAV 536
DB	484	PGIGK	KEKPA	DFKPLGV	KVDVQDP-----ARGTKQLDKEAGLANVWKEQSVLL 536
QY	537	TDTT	TFRDA	HQSLLAT	RVRSFALKPAEAVAKLTPELLSVEAWGATVDVAMRFTFEDPMD 596
DB	537	TDTT	TFRDA	HQSLLAT	RISRDLKKIANTPAAIWPELFSMENNMGATFEDVAYRFLKEDPWK 596
QY	597	RLDEL	REAMP	VNVTQ	MLLRGENTVGYTPYDPSVCRA FVKEAASGSDIFRIFDALNDYSQ 656
DB	597	RLDEL	REVPNTL	QMLLR	SSNANGVYTNYPDNVIKETVKSQAQSGIDVFRIFDSLNWVK 656
QY	657	MRPAID	AVIET	NTTAA	VEAMAYSGLDSPNEKLYTLDYLRKVAEEIVKSGAHILAIKDMA 716

```

Db      657 MTALDAVDTG-KVAEAAICVTGDIIDKNNRTKYDAYTSWAKLEAAAGAHILGIKMA 715
Qy      717 GLLRPAATVLTALRREFDLPVHVHTHTDAGQQLATYFAAAQAGADAVDASAPLSGTT 776
Db      716 GLLKPOAAVELVSALKETIDIPVHLTHDTSNGIYMYAKAVEAGVDIIDVAVSSMAGLT 775
Qy      777 SPSLSAIVAAPHTRTDGLSLEAVSDLEPYKEAVGLYLPESCTPGTPTGRVYHEHP 836
Db      776 SPSASGFYHAMEGNDRPMNVGGVELLSQYWEVKYIYSEFESGKSPHTEIYEHP 835
Qy      837 GGQSLNRAQATLGLADRELIEDNYAANENLGRPTKVTPSSKVVGDALHLVAGVD 896
Db      836 GGQYSLNQQAAGVGLGRNVEKMYRRVNDMGDIKVPKSSKVVGDALVYVQNNLT 895
Qy      897 PAPAADPOKYDIPDSVIAFLRGLNPPGOWPEPLRTRALEGRSGKAPLIEVPEE--- 953
Db      896 EKDYVEGESLDPPDSVVELFGNIGQPHGFFPEKLQKLLKQE-----PITVRPGLLE 951
Qy      954 -----EQAHLDADDKSKERNNSLNRLLFPXPTBEFFLEHRRRRFNGTSALDREFF 1001
Db      952 PVSPEAIKQEFKEQHNLEISD---QDAVAYALVPKVTDYVKTTSYGDISVLDTPTEF 1007
Qy      1002 YGLVEGRETILRLPDVTPLLRLDALISEPDDKGMNVVANVNGOIRPWRVDRSVESVT 1061
Db      1008 YGMTLGEIEIYERKGT-LIVKLISIGEPQDTRVVFELNGQPREVVIKDESIXSV 1066
Qy      1062 ATAEBKADSSKNHGAAPFAGVVTVAE-GDEVKAGDAVAIIIEAMKWEATITASVDGKID 1120
Db      1067 QERLKARTNPISHAAMPGTIVIKVLAEGAKYKNGDHLNEMAKMVEITVQAPPSGTIK 1126
Qy      1121 RVVPAATKVEGGDLIVV 1139
Db      1127 QVHVNGEPIQTGDLLEI 1145

```

RESULT 14

```

US-10-282-122A-60488
; Sequence 60488, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chisen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

```

```

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60488
; LENGTH: 1146
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
; US-10-282-122A-60488

Query Match      43.1%; Score 2493; DB 12; Length 1146;
Best Local Similarity 45.8%; Pred. No. 1.3e-187;
Matches 526; Conservative 185; Mismatches 408; Indels 30; Gaps 11;

Qy      13 KKIIVANRGETIATVAFPAALAEETGAATVAIYPRDRSGSFHRSFASAVRIGTEGSPVKAYL 72
Db      5 KKVIVANRGETIATVAFPAALAEETGAATVAIYPRDRSGSFHRSFASAVRIGTEGSPVKAYL 64
Qy      73 DIDEIIGAAGKAKVADAIYPGYFLSENAQLARECAENGIITFIPTPEVLDLTGDKSAVT 132
Db      65 DIENIEIAKESGADAIHPGVGFLSENIETARRCEQEGIIIFVGPXSKHLDMPGDKIKAKE 124
Qy      133 AAKKAGIPLVAEST-PSKNIDEIVKSAEGOTYPIFKAVAGGGGRGMRVFASPDLEKLA 191
Db      125 QALLADIPVIPSGNPGVAGIKEVEEFGKNGKGYPLMIKASLGCGGGRGMRVVESEKHEKESF 184
Qy      192 TEASREAAAPGDGAVVVERAVINPOHIEVQILGDHTGEVVHLYERDCSLQRHKKVVEI 251
Db      185 ERASSEAKAAGNDEVVVEKVMWPKHIEVQILGDHTGNIVHLPDRDCSLQRHKKVVEI 244
Qy      252 APAQHLDELDRICADAVKRCRISGICGAGTVEFLVDEKGNHVFIEKNPQIOVEHTVTE 311
Db      245 APCNAITSELNRICDAAVKLMKNVDVINAQTVFELV-EGDDVFIEVNPVQVEHTEITE 303
Qy      312 EVTEVDLVKAQMLAAGATLXELGIT---QDKITHGAALQCRITTEDPNNNGFRPDTGTI 368
Db      304 MTGIDIVQSOLFADGVALHDQLVAIPKQBDIHGSAIQSRITTEDPLNFMPTDGRV 363
Qy      369 TAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVMTCRGSDPETAVARAQALAEFTVSG 427
Db      364 DTYRSTGGVRLDAGNGFQGTVVTFFYDSLVLKLTGWTTFEQATKORRNLEFFIRG 423
Qy      428 VATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDADVTVNKEHG 487
Db      424 VKTNPFLNVVRHDPDFASGNVNTSFIDTTPELKFFPHIRDRGKTLRYIGNVTVNGPFG 483
Qy      488 VRPKDVAAPIDKLPNIKOLP-----LPRGSRRLKQLGPAAPFARDLEQDALAVDTTDFR 543
Db      484 IKHRD--KPVVAAEPLRPKIPYQSISPGTKQILDAGPEGVVDVWKQEEVLLDITLIRD 541
Qy      544 AHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAGGATYDVAMRFLPEDPWDLRLDELRE 603
Db      542 AHQSLLATRVRSKQIFQVADAWAHLPMNFSEFMWGGATFDVAYRFLNEDPWRLETLEK 601
Qy      604 AMPNVNIQMLLRGNTVGYTTPYDSCVAFKVEAASSGVDFIRFDALNDVYSQVRPAIDA 663
Db      602 QIPNVMFQMLLAGANAVGVKNYVIREFVKQSAQSGVDVFRVFDLSNLWIKGMEVSIDA 661
Qy      664 VLETNVTAVEAWAYSGDLSQNEKLYTLDDYKMAEIEVKSGAHILAIKDMAGLLRFAA 723
Db      662 VREAG-KVVEAAICVTGDIIDDDTRKYTIIDYKMAKELVAGGTHILGDKMAGLLKQQA 720
Qy      724 VTKLVTLARREFDLPVHVHTHTDAGQQLATYFAAAQAGADAVDASAPLSGTSFSLA 783
Db      721 AVRELIGELKDTVDVPIHLHTDTSNGIYVAAAASAGVDIVDVASSAMSGATSFQSMTG 780
Qy      784 IVAAFAHTRTDGLSLEAVSDLEPYKEAVGLYLPESCTPGTPTGRVYHEHPGQSLNL 843
Db      781 LYYGLVNGRQTNLDAQNSQIINHMYEDVRHYHKDFDNALNSPQTEVITHMPPGQYINL 840

```


QY	1014	LPDVRTPLLVRLDAISEPDDKGMENVTANVNGQIRPMRVDRDRSVESVTATAEK---	ADSS	1070
Db	1015	IAEGKT-LWQQLHIGKLDKQGNFTLVFEVNGNRREIKIKDKVSTKSEIVEEIVADSS		1073
QY	1071	NKGHVAAPPFAG-VVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVPAAK		1129
Db	1074	NKEIGASIPGNVVKVFKPGDKYKKGDSLMLVIEAMKMEINVSVEDGTGIGIFVKEGDQ		1133
QY	1130	VEGGDLIV		1137
Db	1134	VQSGQLLV		1141

Search completed: March 24, 2004, 22:50:25
Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:36:06 ; Search time 64 seconds
(without alignments)
5032.877 Million cell updates/sec

Title: US-10-045-072-2

Perfect score: 5788

Sequence: 1 MSTHTSTLPAPKILVANR.....RVVPAATKVBGGDLIVVVS 1140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: A_Geneseq_29Jan04:*
 - 2: Geneseqp1980s:*
 - 3: Geneseqp1990s:*
 - 4: Geneseqp2000s:*
 - 5: Geneseqp2000s:*
 - 6: Geneseqp2002s:*
 - 7: Geneseqp2003as:*
 - 8: Geneseqp2003bs:*
 - 9: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5788	100.0	1140	4	AAB67129 Coryneb
2	5788	100.0	1140	4	AG90511 C glutami
3	5788	100.0	1140	5	AAB25601 Coryneb
4	5788	100.0	1140	5	AU98053 Coryneb
5	5788	100.0	1140	7	ABU10426 Coryneb
6	5784	99.9	1141	3	ABU10426 Pyruvate
7	5782	99.9	1140	2	AAB93371 C. glutam
8	5780	99.9	1140	4	AG93349 C glutami
9	5759	99.5	1140	5	AU98052 Coryneb
10	5759	99.5	1157	5	AU98050 Coryneb
11	5302.5	91.6	1139	4	AAB83180 Coryneb
12	4647.5	80.3	1141	6	ABU25961 Protein e
13	3791.5	65.5	1127	6	ABU34026 Protein e
14	3712.5	64.1	1127	6	ABU36886 Protein e
15	3712.5	64.1	1127	6	ABU34026 Protein e
16	3483	60.2	1124	7	ADB74461 Mycobacte
17	2621	45.3	532	4	AAB79302 Coryneb
18	2552	44.1	1144	6	ABU24853 Protein e
19	2526	43.6	1148	6	ABU18942 Protein e
20	2525.5	43.6	1178	7	ADB62415 Rat Prote
21	2525.5	43.6	1178	7	ADB62419 Rat Prote
22	2524	43.6	1148	4	AU000511 Bacillus
23	2520.5	43.5	1178	7	ADB62421 Human Pro
24	2520.5	43.5	1178	7	ADB62417 Human Pro
25	2493	43.1	1146	5	ABB47612 Listeria

26	2493	43.1	1146	6	ABU32564	Abu32564 Protein e
27	2491.5	43.0	1144	6	ABU24164	Abu24164 Protein e
28	2488	43.0	1147	4	AAU33972	AAU33972 Staphyloc
29	2488	43.0	1150	6	ABU16467	Abu16467 Protein e
30	2485	42.9	1156	6	ABM73249	Abm73249 Staphyloc
31	2472.5	42.7	1142	4	AAU35213	AAU35213 Enterococ
32	2464.5	42.6	1142	6	ABU30018	Abu30018 Protein e
33	2464	42.6	1151	4	AAAG82677	AAg82677 S. epider
34	2464	42.6	1154	5	ABP38583	Abp38583 Staphyloc
35	2463.5	42.6	1142	6	ABU14604	Abu14604 Protein e
36	2455	42.4	1144	6	ADB08196	ADB08196 Alloiococ
37	2441	42.2	1181	4	AB866604	AB866604 Drosophil
38	2441	42.2	1181	4	AB866605	AB866605 Drosophil
39	2441	42.2	1181	4	AB858211	AB858211 Drosophil
40	2433.5	42.0	1196	4	AB867309	AB867309 Drosophil
41	2405	41.6	1136	5	AB853980	AB853980 Lactococ
42	2375	41.0	1143	6	ABU24993	Abu24993 Protein e
43	2357	40.7	461	4	AA879303	AA879303 Coryneb
44	2299	39.7	1073	4	AAU36788	AAU36788 Staphyloc
45	2174	37.6	973	7	ADC95183	ADC95183 E. faeciu

ALIGNMENTS

RESULT 1

AA867129 AAB67129 standard; protein; 1140 AA.

XX AAB67129;

DT 12-APR-2001 (first entry)

XX Corynebacterium glutamicum pyruvate carboxylase.

XX Pyruvate carboxylase; anaplerotic pathway; industrial fermentation;
XX oxaloacetate.

XX Corynebacterium glutamicum.

OS US6171833-B1.

PN 09-JAN-2001.

XX 23-DEC-1998; 98US-00220081.

XX 23-DEC-1998; 98US-00220081.

XX (MASI) MASSACHUSETTS INST TECHNOLOGY.

FA Sinskey AJ, Lessard PA, Willis LB;

XX WPI: 2001-122330/13.

DR N-PSDB; AAF32185.

XX Novel nucleic acid encoding pyruvate carboxylase from Corynebacterium
XX glutamicum, for replenishing oxaloacetate consumed during lysine and
XX glutamic acid production in industrial fermentations.

PS Claim 1; Col 31-36; 29pp; English.

XX The present invention provides the protein and coding sequences of the
XX Corynebacterium glutamicum pyruvate carboxylase protein. This is an
XX enzyme in the anaplerotic pathway. It can be used in the replenishment of
XX oxaloacetate consumed during lysine and glutamic acid production in
XX industrial fermentation

SQ Sequence 1140 AA;

Query Match 100.0%; Score 5788; DB 4; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

app 1 con

Qy	1	MSHTSTSLT	PAFKKILVAN	GEI	IAV	RAFR	AA	LT	GA	T	GA	T	VA	I	Y	P	R	D	R	G	S	F	H	R	S	F	A	S	E	A	V	R	60		
Db	1	MSHTSTSLT	PAFKKILVAN	GEI	IAV	RAFR	AA	LT	GA	T	GA	T	VA	I	Y	P	R	D	R	G	S	F	H	R	S	F	A	S	E	A	V	R	60		
Qy	61	IGTEGSPVK	AYL	D	I	E	I	G	A	K	K	V	K	A	D	A	I	Y	P	G	G	F	L	S	E	N	A	Q	A	R	E	A	120		
Db	61	IGTEGSPVK	AYL	D	I	E	I	G	A	K	K	V	K	A	D	A	I	Y	P	G	G	F	L	S	E	N	A	Q	A	R	E	A	120		
Qy	121	LDTITGDKS	RAV	T	A	A	K	K	A	G	L	P	V	L	A	B	S	T	P	S	K	N	I	D	I	V	K	S	A	B	E	O	T	180	
Db	121	LDTITGDKS	RAV	T	A	A	K	K	A	G	L	P	V	L	A	B	S	T	P	S	K	N	I	D	I	V	K	S	A	B	E	O	T	180	
Qy	181	VASPD	EL	R	K	L	A	T	E	A	S	R	E	A	E	A	F	G	D	G	A	Y	V	E	R	A	V	I	N	P	Q	H	I	240	
Db	181	VASPD	EL	R	K	L	A	T	E	A	S	R	E	A	E	A	F	G	D	G	A	Y	V	E	R	A	V	I	N	P	Q	H	I	240	
Qy	241	LQRRH	Q	K	V	E	I	A	P	A	O	H	L	D	P	E	L	R	O	R	I	C	A	D	A	V	K	F	C	R	S	I	G	Y	300
Db	241	LQRRH	Q	K	V	E	I	A	P	A	O	H	L	D	P	E	L	R	O	R	I	C	A	D	A	V	K	F	C	R	S	I	G	Y	300
Qy	301	PRIQ	VE	T	V	E	E	V	D	L	V	K	A	Q	M	L	A	G	A	T	L	K	E	L	G	L	T	Q	D	K	I	K	H	360	
Db	301	PRIQ	VE	T	V	E	E	V	D	L	V	K	A	Q	M	L	A	G	A	T	L	K	E	L	G	L	T	Q	D	K	I	K	H	360	
Qy	361	FRD	T	G	I	T	A	Y	R	S	P	G	G	A	G	V	R	D	G	A	O	L	G	G	E	I	T	A	H	F	S	D	M	420	
Db	361	FRD	T	G	I	T	A	Y	R	S	P	G	G	A	G	V	R	D	G	A	O	L	G	G	E	I	T	A	H	F	S	D	M	420	
Qy	421	AEPT	V	S	G	V	A	T	N	I	G	F	L	R	A	L	L	E	E	D	T	S	K	R	I	A	T	G	I	A	D	H	P	480	
Db	421	AEPT	V	S	G	V	A	T	N	I	G	F	L	R	A	L	L	E	E	D	T	S	K	R	I	A	T	G	I	A	D	H	P	480	
Qy	481	TVN	K	P	H	V	R	P	K	D	V	A	A	P	I	D	K	L	P	N	I	K	D	L	P	L	P	R	G	S	R	D	L	540	
Db	481	TVN	K	P	H	V	R	P	K	D	V	A	A	P	I	D	K	L	P	N	I	K	D	L	P	L	P	R	G	S	R	D	L	540	
Qy	541	FRD	A	H	Q	S	L	A	T	R	V	R	S	F	A	K	P	A	A	E	A	V	A	K	I	T	E	L	S	V	E	A	W	600	
Db	541	FRD	A	H	Q	S	L	A	T	R	V	R	S	F	A	K	P	A	A	E	A	V	A	K	I	T	E	L	S	V	E	A	W	600	
Qy	601	L	R	E	A	M	P	N	V	N	I	Q	M	L	L	R	G	N	T	V	G	T	P	Y	P	D	S	V	C	R	A	F	V	660	
Db	601	L	R	E	A	M	P	N	V	N	I																								

Db 1081 GVTVTVAGDEVKAGDAVAIIEMKQKEATITASVDGKIDRVVVPAATKVEGGDLIVWS 1140

RESULT 2
AAG90511
XX ID AAG90511 standard; protein; 1140 AA.
XX AC AAG90511;
XX XX
XX 26-SBP-2001 (first entry)
XX C glutamicum protein fragment SEQ ID NO: 4265.
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX KW
XX OS Corynebacterium glutamicum.
XX EF1108790-A2.
XX XX
XX 20-JUN-2001.
XX XX
XX 18-DEC-2000; 2000EP-00127688.
XX PR 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-0015162.
XX PR 03-AUG-2000; 2000JP-00280988.
XX XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX XX
XX WPI; 2001-376931/40.
XX DR N-FSDB; AAH65730.
XX XX
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 17; SEQ ID NO 4265; 245pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
XX sequences from the Corynebacterium Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of corynebacterium, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX Corynebacterium, and identifying a homologue of a gene derived from
XX Corynebacterium. Corynebacterium bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a protein described in the
XX exemplification of the invention. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from the European Patent Office
XX
XX Sequence 1140 AA;
XX
XX Query Match 100.0%; Score 5788; DB 4; Length 1140;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MSTHTSSTLPAPFKKILVANGETIAVRAAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
Db 1 MSTHTSSTLPAPFKKILVANGETIAVRAAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
Oy 61 IGTEGSPVKAYLDIDEIIIGAARKVKADAIYPGYGFLSENAQLARECAENGITTFIGTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIIGAARKVKADAIYPGYGFLSENAQLARECAENGITTFIGTPEV 120
Oy 121 LDLTGDKSRVATAAKAGLPVLAESTPSKNIDEIVKSABGQTYIFVKAVAGGGGRNRF 180
Db 121 LDLTGDKSRVATAAKAGLPVLAESTPSKNIDEIVKSABGQTYIFVKAVAGGGGRNRF 180

QY 421 AEFTVSGVATNIGFRLALLREEDFTSKHATGFIADPHLLQAPPADEQGRILDLADY 480
 Db |||||
 QY 421 AEFTVSGVATNIGFRLALLREEDFTSKHATGFIADPHLLQAPPADEQGRILDLADY 480
 Db |||||
 QY 481 TVNKGHGVKPDVAAPIDKLNKIDPLPRGSRDLKQGPAAFPARDLRQDALAVDTT 540
 Db |||||
 QY 481 TVNKGHGVKPDVAAPIDKLNKIDPLPRGSRDLKQGPAAFPARDLRQDALAVDTT 540
 Db |||||
 QY 541 FRDAHQSLLATNRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPMDRLDE 600
 Db |||||
 QY 541 FRDAHQSLLATNRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPMDRLDE 600
 Db |||||
 QY 601 LREAMPNVIQMLRGRNTGTYTPYDSVCRAFVKEAASSGVDIFRFDALNDVSQRP 660
 Db |||||
 QY 601 LREAMPNVIQMLRGRNTGTYTPYDSVCRAFVKEAASSGVDIFRFDALNDVSQRP 660
 Db |||||
 QY 661 IDAVLETTNVAEAVAMAYSGDLSDBNEKLYTLDYLVKVAEIVKSGAHILAIDKMAGLR 720
 Db |||||
 QY 661 IDAVLETTNVAEAVAMAYSGDLSDBNEKLYTLDYLVKVAEIVKSGAHILAIDKMAGLR 720
 Db |||||
 QY 721 PAAVTKLVTLRREPDLPHVHTHTDAGGQLATYFAAAQAGADAVDGASAPLSGTTSPS 780
 Db |||||
 QY 721 PAAVTKLVTLRREPDLPHVHTHTDAGGQLATYFAAAQAGADAVDGASAPLSGTTSPS 780
 Db |||||
 QY 781 LSAIIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRVYRHEIPGQL 840
 Db |||||
 QY 781 LSAIIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRVYRHEIPGQL 840
 Db |||||
 QY 841 SNLRAQATAGLADRFELIEDNYAAVNEMLGRPKVTTPSSKVVGDALHLVAGVDPADF 900
 Db |||||
 QY 841 SNLRAQATAGLADRFELIEDNYAAVNEMLGRPKVTTPSSKVVGDALHLVAGVDPADF 900
 Db |||||
 QY 901 AADPKYDIPDSVIAFLRGEIENPPGWPBPLRTRALEGRSEGKAPLTPVEPEEQALDA 960
 Db |||||
 QY 901 AADPKYDIPDSVIAFLRGEIENPPGWPBPLRTRALEGRSEGKAPLTPVEPEEQALDA 960
 Db |||||
 QY 961 DDSKERRSNLRLLFPKTEEFLEHRRRFGNTSALDDREFFYGLVGEKRETLIRLPDVRTP 1020
 Db |||||
 QY 961 DDSKERRSNLRLLFPKTEEFLEHRRRFGNTSALDDREFFYGLVGEKRETLIRLPDVRTP 1020
 Db |||||
 QY 1021 LLVLRLDAISEDDKGRNVANVNGQIRPMRVDRSRESVTATAEKADSSNKGHVAAPFA 1080
 Db |||||
 QY 1021 LLVLRLDAISEDDKGRNVANVNGQIRPMRVDRSRESVTATAEKADSSNKGHVAAPFA 1080
 Db |||||
 QY 1081 GVVTVVAGDEBKAGDAVAIEAMKMEATITASVDGKIDRVVVPAAKTVGEGDLIVVVS 1140
 Db |||||
 QY 1081 GVVTVVAGDEBKAGDAVAIEAMKMEATITASVDGKIDRVVVPAAKTVGEGDLIVVVS 1140
 Db |||||

RESULT 4
 ID AAU98053
 AC AAU98053
 AC AAU98053;
 DT 27-AUG-2002 (first entry)
 DE Corynebacterium wild-type feedback-resistant pyruvate carboxylase enzyme.
 KW Feedback-resistant; pyruvate carboxylase; enzyme;
 KW aspartic acid feedback inhibition resistant.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN WO200231158-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 12-OCT-2001; 2001WO-US031893.
 XX
 PR 13-OCT-2000; 2000US-0239913P.
 XX

(ARCH) ARCHER-DANIELS MIDLAND CO.
 Hanke PD;
 WPI; 2002-463267/49.
 Novel mutated, feedback resistant pyruvate carboxylase enzyme
 polypeptide, useful for producing amino acids e.g. L-lysine, L-threonine,
 L-glycine, L-glutamic acid, L-proline and L-methionine and L-isoleucine.
 Disclosure; Fig 2; 42pp; English.
 The present invention relates to a new mutated, feedback-resistant
 pyruvate carboxylase enzyme. The invention is useful for producing an
 amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Arg and L-Pro), by
 culturing a host cell in a suitable media and separating the amino acid
 from the medium. The vector of the invention is useful for replacement of
 a wild-type pyruvate carboxylase gene, with a feedback resistant pyruvate
 carboxylase gene, in a Corynebacterium glutamicum, by replacing a genomic
 copy of the wild-type pyruvate carboxylase gene with a selectable marker
 gene through homologous recombination to form a first recombinant
 strain, and replacing the selectable marker gene in the first recombinant
 strain, with feedback resistant pyruvate carboxylase gene through
 homologous recombination to form a second recombinant strain, where the
 homologous recombination in the above steps, occurs between the host cell
 and the vector. The feedback-resistant pyruvate carboxylase enzyme is
 resistant to feedback inhibition from aspartic acid. The present amino
 acid sequence represents the wild-type feedback-resistant pyruvate
 carboxylase enzyme of the invention
 Sequence 1140 AA;
 Query Match 100.0%; Score 5788; DB 5; Length 1140;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSTHTSSTLPAPFKILVANRGEIAVRAFAALLETGAATVAIYPREDRGSFHRSFASEAVR 60
 Db 1 MSTHTSSTLPAPFKILVANRGEIAVRAFAALLETGAATVAIYPREDRGSFHRSFASEAVR 60
 QY 61 IGTEGSPVAYLDDIEITGAANKVADAIYPGYFLSENAQLARECAENGITFIPTPEV 120
 Db 61 IGTEGSPVAYLDDIEITGAANKVADAIYPGYFLSENAQLARECAENGITFIPTPEV 120
 QY 121 LDLTGDKSAVTAATAKAGLPVLAESTPSKNIDIEIVKSAEGQTYPIFVKAVAGGGGRGWR 180
 Db 121 LDLTGDKSAVTAATAKAGLPVLAESTPSKNIDIEIVKSAEGQTYPIFVKAVAGGGGRGWR 180
 QY 181 VASPDRLKRLATEASREAEAFDGAUVVERAVINPQHIIEVQILGDHTGEVHLVERDCS 240
 Db 181 VASPDRLKRLATEASREAEAFDGAUVVERAVINPQHIIEVQILGDHTGEVHLVERDCS 240
 QY 241 LORRHOKVVEIAPACHLDPEDLRICADAVKFCRIGYQGAGTVFVLDEKGNHVFIE 300
 Db 241 LORRHOKVVEIAPACHLDPEDLRICADAVKFCRIGYQGAGTVFVLDEKGNHVFIE 300
 QY 301 PRIQVEHTVTEVTEVDLVKQMLAAGATLKELGLTQDKINHTGAALQCRITTTEDP 360
 Db 301 PRIQVEHTVTEVTEVDLVKQMLAAGATLKELGLTQDKINHTGAALQCRITTTEDP 360
 QY 361 FRPDGTITTAIYRSPGAGVRLDGAALGGEITTAHFDLSMLVKMTCRGSDFFETAVARAQAL 420
 Db 361 FRPDGTITTAIYRSPGAGVRLDGAALGGEITTAHFDLSMLVKMTCRGSDFFETAVARAQAL 420
 QY 421 AEFTVSGVATNIGFRLALLREEDFTSKHATGFIADPHLLQAPPADEQGRILDLADY 480
 Db 421 AEFTVSGVATNIGFRLALLREEDFTSKHATGFIADPHLLQAPPADEQGRILDLADY 480
 QY 481 TVNKGHGVKPDVAAPIDKLNKIDPLPRGSRDLKQGPAAFPARDLRQDALAVDTT 540
 Db 481 TVNKGHGVKPDVAAPIDKLNKIDPLPRGSRDLKQGPAAFPARDLRQDALAVDTT 540
 QY 541 FRDAHQSLLATNRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPMDRLDE 600

Db 541 FDAHQSLATVRGKALPAAEAVALTPPELLSVEANGGATYVAVRFLFEDPWRDLDE 600
 QY 601 LREAMPNVIQMLGRNTVGTTPYDPSVCRAVFEAAASSGVDPRIIFDALNDVSQMRPA 660
 Db 601 LREAMPNVIQMLGRNTVGTTPYDPSVCRAVFEAAASSGVDPRIIFDALNDVSQMRPA 660
 QY 661 IDAVLENTNVAEAVAMAYSGDLSDNKELYLDYLLKWAEEIVKSGAHILAIDKMAGLLR 720
 Db 661 IDAVLENTNVAEAVAMAYSGDLSDNKELYLDYLLKWAEEIVKSGAHILAIDKMAGLLR 720
 QY 721 PAAVTKLVTLRRREFDLPVHVHTHTAGGQATYFAAAQAGADAVDGSAPLSGTTSQPS 780
 Db 721 PAAVTKLVTLRRREFDLPVHVHTHTAGGQATYFAAAQAGADAVDGSAPLSGTTSQPS 780
 QY 781 LSAIVAAFAHTRDRDGLSLAEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
 Db 781 LSAIVAAFAHTRDRDGLSLAEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
 QY 841 SNLRAQATAGLADRPEDLIEDNYAAVNEMLGRPTKVTPTSSKVVGDALHLVAGVDPADP 900
 Db 841 SNLRAQATAGLADRPEDLIEDNYAAVNEMLGRPTKVTPTSSKVVGDALHLVAGVDPADP 900
 QY 901 AADPOKYDIPDSVIAFLRGELNPGGWPPEPLRTRALGRSEKAPLFEVPEEEOAHLDA 960
 Db 901 AADPOKYDIPDSVIAFLRGELNPGGWPPEPLRTRALGRSEKAPLFEVPEEEOAHLDA 960
 QY 961 DSKERNNSLNLLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVGEGETLIRLPDVRT 1020
 Db 961 DSKERNNSLNLLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVGEGETLIRLPDVRT 1020
 QY 1021 LLAIVLDAISEPDDKGNVNVANVNGOIRPMRVDRSVSVTATAEKADSSNKGHVAAEPA 1080
 Db 1021 LLAIVLDAISEPDDKGNVNVANVNGOIRPMRVDRSVSVTATAEKADSSNKGHVAAEPA 1080
 QY 1081 GVVTVTVASGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140
 Db 1081 GVVTVTVASGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140

RESULT 5
 ABU10426
 ID ABU10426 standard; protein; 1140 AA.
 XX
 AC ABU10426;
 XX
 DT 25-AUG-2003 (first entry)
 XX
 DE Corynebacterium glutamicum pyruvate carboxylase.
 XX
 KW Pyruvate carboxylase; gene; anaplerotic enzyme; oxaloacetate;
 KW biosynthesis; growth; lysine production; glutamic acid production;
 KW industrial fermentation; enzyme.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN US2003027305-A1.
 XX
 PD 06-FEB-2003.
 XX
 PF 15-JAN-2002; 2002US-00045072.
 XX
 PR 23-DEC-1998; 98US-00220081.
 PR 03-OCT-2000; 2000US-00677575.
 XX
 PA (ARCH) ARCHER-DANIELS MIDLAND CO.
 XX
 PI Sinskey AJ, Lessard PA, Willis LB;
 XX
 DR WPI; 2003-479542/58.
 DR N-PSDB; ACA62133.
 XX
 PT New pyruvate carboxylase from Corynebacterium glutamicum, useful as an

PT anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis
 PT during growth, or for lysine or glutamic acid production in industrial
 PT fermentations.
 XX Claim 1; Page 16-19; 29pp; English.
 PS
 XX The invention describes a new isolated pyruvate carboxylase polypeptide
 CC having an amino acid sequence at least 95% identical to a sequence
 CC comprising 1140 amino acids from Corynebacterium glutamicum, or the
 CC complete amino acid sequence encoded by the cosmid clone deposited with
 CC the American Type Culture Collection. The polypeptide is useful as an
 CC anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis
 CC during growth. The polypeptide is also useful for lysine or glutamic acid
 CC production in industrial fermentations. This is the amino acid sequence
 CC of Corynebacterium glutamicum pyruvate carboxylase
 XX
 SQ Sequence 1140 AA;
 Query Match 100.0%; Score 5788; DB 7; Length 1140;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSTHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAVR 60
 Db 1 MSTHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAVR 60
 QY 61 IGTEGSPVKAYLDIDIIIGAAKKVKADAIYPGVLSENAQLARECAENGITFIPTPEV 120
 Db 61 IGTEGSPVKAYLDIDIIIGAAKKVKADAIYPGVLSENAQLARECAENGITFIPTPEV 120
 QY 121 LDITGKSRAVTAAKKAGLPVLAESTPSKNIIDIVKSAEGQTYPIFVKAVAGGGGGRMF 180
 Db 121 LDITGKSRAVTAAKKAGLPVLAESTPSKNIIDIVKSAEGQTYPIFVKAVAGGGGGRMF 180
 QY 181 VASDELKRLATASAEAEAAFGGAVYVERAVINPQIHIEVQILGHTGEVHLIERDCS 240
 Db 181 VASDELKRLATASAEAEAAFGGAVYVERAVINPQIHIEVQILGHTGEVHLIERDCS 240
 QY 241 LQRRHQKVEIAPAQHLDPDLRDIICADAVKFCRSIGYCGAGVFEFLVDEKGNHVIEMN 300
 Db 241 LQRRHQKVEIAPAQHLDPDLRDIICADAVKFCRSIGYCGAGVFEFLVDEKGNHVIEMN 300
 QY 301 PRIQVEHTVTEVTEVDLVKAAQRLAAGATLKELGTLQDKIKTHGAALQCRITTEPNNG 360
 Db 301 PRIQVEHTVTEVTEVDLVKAAQRLAAGATLKELGTLQDKIKTHGAALQCRITTEPNNG 360
 QY 361 FRDPTGTITAYRSPGGAGVRLDCAAOIIGBEITAHFDSMLVKMTCRSDPETAVARQAL 420
 Db 361 FRDPTGTITAYRSPGGAGVRLDCAAOIIGBEITAHFDSMLVKMTCRSDPETAVARQAL 420
 QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADV 480
 Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADV 480
 QY 481 TVNKHGVRPKDVAAPIDKLPNIKDLPLPRGSDRLKQLGPAFAFADRLREODALAVTDTT 540
 Db 481 TVNKHGVRPKDVAAPIDKLPNIKDLPLPRGSDRLKQLGPAFAFADRLREODALAVTDTT 540
 QY 541 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEANGGATYDVAMRFLFEDPWRDLDE 600
 Db 541 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEANGGATYDVAMRFLFEDPWRDLDE 600
 QY 601 LREAMPNVIQMLGRNTVGTTPYDPSVCRAVFEAAASSGVDPRIIFDALNDVSQMRPA 660
 Db 601 LREAMPNVIQMLGRNTVGTTPYDPSVCRAVFEAAASSGVDPRIIFDALNDVSQMRPA 660
 QY 661 IDAVLENTNVAEAVAMAYSGDLSDNKELYLDYLLKWAEEIVKSGAHILAIDKMAGLLR 720
 Db 661 IDAVLENTNVAEAVAMAYSGDLSDNKELYLDYLLKWAEEIVKSGAHILAIDKMAGLLR 720
 QY 721 PAAVTKLVTLRRREFDLPVHVHTHTAGGQATYFAAAQAGADAVDGSAPLSGTTSQPS 780
 Db 721 PAAVTKLVTLRRREFDLPVHVHTHTAGGQATYFAAAQAGADAVDGSAPLSGTTSQPS 780

Db 1022 LVLRLDAISEPDGKRNVRVAVNVNQIRPMVRDRSVSTATAEKADSSNKGHVAAFFA 1081
QY 1081 GVVTVTVAGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140
Db 1082 GVVTVTVAGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1141
RESULT 7
ID AAW93971 standard; protein; 1140 AA.
AC AAW93971;
XX
DT 30-JUN-1999 (first entry)
XX
DE C. glutamicum pyruvate carboxylase protein.
XX
KW Pyruvate carboxylase; amino acid production; lysine production;
KW threonine production; homoserine production; glutamate production;
KW arginine production; feed additive; condiment; pharmaceutical;
KW fine chemical; ss.
XX
XX Corynebacterium glutamicum.
XX
XX DE19831609-Al.
XX
XX 15-APR-1999. 98DE-01031609.
XX
XX 14-JUL-1998;
XX
XX 04-OCT-1997; 97DE-01043894.
PR (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
XX Peters-Wendisch P, Eikmanns B, Sahm H;
XX
XX WPI; 1999-245521/21.
DR N-PSDB; AAX24102.
XX
PT Increasing microbial production of specific amino acids by increasing
PT activity or expression of pyruvate carboxylase.
XX
XX Claim 15; Page 11-15; 18pp; German.
XX
XX This invention describes the isolation of a pyruvate carboxylase from
CC Corynebacterium glutamicum which is used in a novel method for production
CC of lysine, threonine, homoserine, glutamate and/or arginine, variously
CC useful as feed additives, condiments, pharmaceuticals and intermediates
CC for fine chemicals. Increasing pyruvate carboxylase activity increases
CC the yield of microbial production of amino acids of the aspartate and/or
CC glutamate families, e.g. about 50% more lysine, 40% more threonine and
CC 150% more homoserine are secreted into the culture medium
XX
XX Sequence 1140 AA;
QY
Query Match 99.9%; Score 5782; DB 2; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSTHTSSTLPAPKKILVANRGEIIVAFRAALETGAATVAIYPREDRGSFHRSEAFVR 60
Db 1 MSTHTSSTLPAPKKILVANRGEIIVAFRAALETGAATVAIYPREDRGSFHRSEAFVR 60
QY 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGVLSENAQLARECAENGITFIGPTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGVLSENAQLARECAENGITFIGPTPEV 120
QY 121 LDITGDKSRVTAAKKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGCGGRMRF 180
Db 121 LDITGDKSRVTAAKKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGCGGRMRF 180
QY 181 VASPELURKLTAEASREAEAFGDGAVYVERAVINPQHTEVQLGDHTGEVHLYERDCS 240

Db 181 VASPELURKLTAEASREAEAFGDGAVYVERAVINPQHTEVQLGDHTGEVHLYERDCS 240
QY 241 LQRRHQKVEIIAPQAHLDPEDLRICADAVKFCRSIGYQAGTVEFLVDEKGNHVEIWN 300
Db 241 LQRRHQKVEIIAPQAHLDPEDLRICADAVKFCRSIGYQAGTVEFLVDEKGNHVEIWN 300
QY 301 PRIQVEHTVTEEVTDLVKAQWRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTDLVKAQWRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
QY 361 PRDPTGTTIAYRSPGGAGVRLDGAAGLGGIITAHFDSMLVKMTCRSDSDTAVARAQAL 420
Db 361 PRDPTGTTIAYRSPGGAGVRLDGAAGLGGIITAHFDSMLVKMTCRSDSDTAVARAQAL 420
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDDQGRILDYADV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDDQGRILDYADV 480
QY 481 TVNKPBGVVRKDVAAPIDKLPNIDLPGRSDRLKOLGPAFAFARDLREODALATDTT 540
Db 481 TVNKPBGVVRKDVAAPIDKLPNIDLPGRSDRLKOLGPAFAFARDLREODALATDTT 540
QY 541 FRDAHQSLLATRVRSFALKEAABAVAKLTPELLSVEAWGATYDVAMRFLEFDPWRLDE 600
Db 541 FRDAHQSLLATRVRSFALKEAABAVAKLTPELLSVEAWGATYDVAMRFLEFDPWRLDE 600
QY 601 LREAMPNNVIOMLLRGNTVGYTPYDPSVCRFAVKEAASGVDFRIFDALNDVSOQRPA 660
Db 601 LREAMPNNVIOMLLRGNTVGYTPYDPSVCRFAVKEAASGVDFRIFDALNDVSOQRPA 660
QY 661 IDAVLENTTAAVEAVAYSGDLSDPNEKLYTLDYILKVAEEIVKSGAHILAIKDMAGLLR 720
Db 661 IDAVLENTTAAVEAVAYSGDLSDPNEKLYTLDYILKVAEEIVKSGAHILAIKDMAGLLR 720
QY 721 PAAVTKLVTLARREFDLPVHVHTHTAGGQALATYFAAAQAGADAVDGSAPLSGTTSSQPS 780
Db 721 PAAVTKLVTLARREFDLPVHVHTHTAGGQALATYFAAAQAGADAVDGSAPLSGTTSSQPS 780
QY 781 LSAITVAFAFTRDRTGSLSEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQOL 840
Db 781 LSAITVAFAFTRDRTGSLSEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQOL 840
QY 841 SNLRAQATALGLADRFELIEDNYAANNMELGRPTKVTTPSSKVVGDALHLVAGVDPAF 900
Db 841 SNLRAQATALGLADRFELIEDNYAANNMELGRPTKVTTPSSKVVGDALHLVAGVDPAF 900
QY 901 AADPKQYDIPDSVIAFLRGLGNPPGPGWPEPLATRALEGRSEKAPLTPVEPEEQHLLDA 960
Db 901 AADPKQYDIPDSVIAFLRGLGNPPGPGWPEPLATRALEGRSEKAPLTPVEPEEQHLLDA 960
QY 961 DDSKERNSNLRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db 961 DDSKERNSNLRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
QY 1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMVRDRSVSTATAEKADSSNKGHVAAFFA 1080
Db 1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMVRDRSVSTATAEKADSSNKGHVAAFFA 1080
QY 1081 GVVTVTVAGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140
Db 1081 GVVTVTVAGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140
RESULT 8
ID AAG93249 standard; protein; 1140 AA.
XX
AC AAG93249;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment mutant P458S.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; mutant; mutein.
 XX Corynebacterium glutamicum.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 458
 FT /note= "wild-type Pro substituted by Ser"
 FT
 XX EP1108790-A2.
 XX 20-JUN-2001.
 XX 18-DEC-2000; 2000EP-00127688.
 XX 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI; 2001-376931/40.
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX Claim 43; Page; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a mutant protein described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX Sequence 1140 AA;
 SQ

Query Match 99.9%; Score 5780; DB 4; Length 1140;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVR 60
 Db 1 MSHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVR 60
 QY 61 IGTGSPVKAYLDIDIEIIICAAKKVADAIYPGYGFLSENAQLARECAENGITFTGPTPEV 120
 Db 61 IGTGSPVKAYLDIDIEIIICAAKKVADAIYPGYGFLSENAQLARECAENGITFTGPTPEV 120
 QY 121 LDLTGDKSRVTAANKAGIPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGWR 180
 Db 121 LDLTGDKSRVTAANKAGIPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGWR 180
 QY 181 VASPELAKLATEASREARAFGDGAUVVERAVINPQHIIEVQILGDHTGEVHHYERDCS 240
 Db 181 VASPELAKLATEASREARAFGDGAUVVERAVINPQHIIEVQILGDHTGEVHHYERDCS 240
 QY 241 LQRRHQKVEIAPACHLPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFTEMN 300
 Db 241 LQRRHQKVEIAPACHLPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFTEMN 300

QY 301 PRIQVEHTVTEEVTEVDLVKQMRLAAGATLKELGLTQDKIKTHGAALOCRTITTEDPNNG 360
 Db 301 PRIQVEHTVTEEVTEVDLVKQMRLAAGATLKELGLTQDKIKTHGAALOCRTITTEDPNNG 360
 QY 361 FRPDTGTITAYRSPGGAGVRLDGAALQGGIITAHFDSMLVKMTCRGSDFETAVAAQRAAL 420
 Db 361 FRPDTGTITAYRSPGGAGVRLDGAALQGGIITAHFDSMLVKMTCRGSDFETAVAAQRAAL 420
 QY 421 AEFTVSGVATNIGFLRALREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
 Db 421 AEFTVSGVATNIGFLRALREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
 QY 481 TVNKPGEVPRKQVAAPIDKLPNIKOLPLPRGSRDLKQLGPAAPARDLREQDALLAVTDTT 540
 Db 481 TVNKPGEVPRKQVAAPIDKLPNIKOLPLPRGSRDLKQLGPAAPARDLREQDALLAVTDTT 540
 QY 541 FRDAHQSLLATVRSPALKAFAEAVAKLTPELLSVEANGGATYDVAMRFLFDPDRDLDE 600
 Db 541 FRDAHQSLLATVRSPALKAFAEAVAKLTPELLSVEANGGATYDVAMRFLFDPDRDLDE 600
 QY 601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCRFAFVKEAASSGVVDIPIFDALNDVSMRPA 660
 Db 601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCRFAFVKEAASSGVVDIPIFDALNDVSMRPA 660
 QY 661 IDAVLEINTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKAAEIVKSGAHLAIKDMAGLLR 720
 Db 661 IDAVLEINTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKAAEIVKSGAHLAIKDMAGLLR 720
 QY 721 PAAVTKLVTALRRFDPVHVHTHTAGQLATYFAAAQAGADAVDGAAPLSGTTSQPS 780
 Db 721 PAAVTKLVTALRRFDPVHVHTHTAGQLATYFAAAQAGADAVDGAAPLSGTTSQPS 780
 QY 781 LSAIVAAFAHTRDTCGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVYRHEIPGGQL 840
 Db 781 LSAIVAAFAHTRDTCGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVYRHEIPGGQL 840
 QY 841 SNLRQAATGALGRPELIEDNYAAVNEMLRPTKVTPTSSKVVGLLHLVAGVDPADP 900
 Db 841 SNLRQAATGALGRPELIEDNYAAVNEMLRPTKVTPTSSKVVGLLHLVAGVDPADP 900
 QY 901 ADPOKQYDPSVIAFLRGLGNPFGWPEPLTRALRSGRSGKAPLIEVPEEQAHLLDA 960
 Db 901 ADPOKQYDPSVIAFLRGLGNPFGWPEPLTRALRSGRSGKAPLIEVPEEQAHLLDA 960
 QY 961 DSKERRNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRLPDVTRP 1020
 Db 961 DSKERRNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRLPDVTRP 1020
 QY 1021 LLVRLDAISEPDDKGRNVVANVNGOIRPMRVDRSVESVTATAKADSSNKGHVAAAPPA 1080
 Db 1021 LLVRLDAISEPDDKGRNVVANVNGOIRPMRVDRSVESVTATAKADSSNKGHVAAAPPA 1080
 QY 1081 GVVTVTVASGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVWPAATKVEGDLIVVVS 1140
 Db 1081 GVVTVTVASGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVWPAATKVEGDLIVVVS 1140

RESULT 9
 AAU98052
 ID AAU98052 standard; protein; 1140 AA.
 XX
 AC AAU98052;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Corynebacterium mutant feedback-resistant pyruvate carboxylase enzyme.
 XX Feedback-resistant; pyruvate carboxylase; enzyme;
 KW aspartic acid feedback inhibition resistant; mutant; mutein.
 XX
 OS Corynebacterium glutamicum.
 OS Synthetic.
 XX

FH	Key	Location/Qualifiers	
FT	Misc-difference 1	/note= "Wild-type Met substituted by Val"	
FT	Misc-difference 153	/note= "Wild-type Glu substituted by Asp"	
FT	Misc-difference 182	/note= "Wild-type Ala substituted by Ser"	
FT	Misc-difference 206	/note= "Wild-type Ala substituted by Ser"	
FT	Misc-difference 227	/note= "Wild-type His substituted by Arg"	
FT	Misc-difference 455	/note= "Wild-type Ala substituted by Gly"	
FT	Region	1110..1122	
FT	Misc-difference 1116	/note= "Specifically claimed in claim 18"	
FT	Misc-difference 1116	/note= "Wild-type Asp substituted by Glu"	
XX	WO200231159-A2.		
XX	18-APR-2002.		
XX	12-OCT-2001; 2001WO-US031893.		
XX	13-OCT-2000; 2000US-0239913P.		
XX	(ARCH) ARCHER-DANIELS MIDLAND CO.		
XX	Hanke PD;		
XX	WPI; 2002-463267/49.		
XX	Novel mutated, feedback resistant pyruvate carboxylase enzyme		
PT	polypeptide, useful for producing amino acids e.g. L-lysine, L-threonine,		
PT	L-glutamine, L-glutamic acid, L-proline and L-methionine and L-isoleucine.		
XX	Claim 1; Page; 42pp; English.		
XX	The present invention relates to a new mutated, feedback-resistant		
CC	pyruvate carboxylase enzyme. The invention is useful for producing an		
CC	amino acid (e.g. L-Lys, L-Thr, L-Ile, L-Glu, L-Arg and L-Pro), by		
CC	culturing a host cell in a suitable media and separating the amino acid		
CC	from the medium. The vector of the invention is useful for replacement of		
CC	a wild-type pyruvate carboxylase gene, with a feedback resistant pyruvate		
CC	carboxylase gene, in a Corynebacterium glutamicum, by replacing a genomic		
CC	copy of the wild-type pyruvate carboxylase gene with a selectable marker		
CC	gene through homologous recombination to form a first recombinant		
CC	strain, and replacing the selectable marker gene in the first recombinant		
CC	strain, with feedback resistant pyruvate carboxylase gene through		
CC	homologous recombination to form a second recombinant strain, where the		
CC	homologous recombination in the above steps, occurs between the host cell		
CC	and the vector. The feedback-resistant pyruvate carboxylase enzyme is		
CC	resistant to feedback inhibition from aspartic acid. The present amino		
CC	acid sequence represents the mutant feedback-resistant pyruvate		
CC	carboxylase enzyme of the invention. Note: The present sequence is not		
CC	shown in the specification but is derived from the wild-type feedback-		
CC	resistant pyruvate carboxylase enzyme (AAU98053) given in figure 2 of the		
CC	specification		
XX	Sequence 1140 AA;		
SQ			
	Query Match	99.5%; Score 5759; DB 5; Length 1140;	
	Best Local Similarity	99.4%; Pred. No. 0;	
	Matches 1133; Conservative	5; Mismatches 2; Indels 0; Gaps 0;	
QY	1	MSHTSTSTPAFKKILVANGETAVRAFAALETGAATVAIYPRDGRGSHFSPASEAVR	60
DB	1	VSTHTSTSTPAFKKILVANGETAVRAFAALETGAATVAIYPRDGRGSHFSPASEAVR	60
QY	61	IGTEGSPVKAYLIDIDIIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIPTPEV	120
DB	61	IGTEGSPVKAYLIDIDIIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIPTPEV	120

RESULT 10
AAU98050
ID AAU98050 standard; protein; 1157 AA.
XX

QY	121	LDLTGDKSRVATRAAKKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGGGGMEF	180
DB	121	LDLTGDKSRVATRAAKKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGGGGMEF	180
QY	181	VASPDRLRLATASREAEAAAFDGDGAVYVERAVINPQHIEVOILGDHTGEVHLIERDCS	240
DB	181	VSSPDRLRLATASREAEAAAFDGDGAVYVERAVINPQHIEVOILGDRTGEVHLIERDCS	240
QY	241	LQRHOKVVEIAPAQHLDPEDRLDRICADAVKCRSIGYQAGTVBEFLVDEKGNHVEIENW	300
DB	241	LQRHOKVVEIAPAQHLDPEDRLDRICADAVKCRSIGYQAGTVBEFLVDEKGNHVEIENW	300
QY	301	PRIOVEHTVEEVTVDLVKAQVRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG	360
DB	301	PRIOVEHTVEEVTVDLVKAQVRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG	360
QY	361	FRPDTGTITAYRSPGGAGVRLDGAALGGEIITAHFDSMLVKMTCRGSDETAVARAQAL	420
DB	361	FRPDTGTITAYRSPGGAGVRLDGAALGGEIITAHFDSMLVKMTCRGSDETAVARAQAL	420
QY	421	AFTVSGVATNIGFLRALLREEDFTSKRIATQFIADPHLLQAPPADDDQGRILDYADV	480
DB	421	AFTVSGVATNIGFLRALLREEDFTSKRIATQFIADPHLLQAPPADDDQGRILDYADV	480
QY	481	TVKNPGRVPRKQVAAPIDKLPNIKPLPRGRDRKQLGPAFAFARDLREQDALAVTDTT	540
DB	481	TVKNPGRVPRKQVAAPIDKLPNIKPLPRGRDRKQLGPAFAFARDLREQDALAVTDTT	540
QY	541	FRDAHOSLLATRVRSFALPAAEAVAKLTPELLSVEANGATYDVAMRFLFEDFPWRLDE	600
DB	541	FRDAHOSLLATRVRSFALPAAEAVAKLTPELLSVEANGATYDVAMRFLFEDFPWRLDE	600
QY	601	LREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFPVKEAASGVVDIFRIFDALNDVSOVRPA	660
DB	601	LREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFPVKEAASGVVDIFRIFDALNDVSOVRPA	660
QY	661	IDAVLETNTAVAVAMAYSGDLSDPNEKLYTLDYILKMAEIVKSGAHILAIKDMAGLLR	720
DB	661	IDAVLETNTAVAVAMAYSGDLSDPNEKLYTLDYILKMAEIVKSGAHILAIKDMAGLLR	720
QY	721	PAAVTKLVTLALRREFDLPVHVHTHTAGGQATYFAAAQAGADAVDGAAPLSGTTSQPS	780
DB	721	PAAVTKLVTLALRREFDLPVHVHTHTAGGQATYFAAAQAGADAVDGAAPLSGTTSQPS	780
QY	781	LSAIVAAFAHTRDTGLSLEAVSDLEPYMEAVGLYLPESGTPGPTGRVYRHEIPGGQL	840
DB	781	LSAIVAAFAHTRDTGLSLEAVSDLEPYMEAVGLYLPESGTPGPTGRVYRHEIPGGQL	840
QY	841	SNLRAQATALGLADRFELIEDNYAAVNEMLGRFTKVTPTSSKVVGDLALHLVGAGVDPADF	900
DB	841	SNLRAQATALGLADRFELIEDNYAAVNEMLGRFTKVTPTSSKVVGDLALHLVGAGVDPADF	900
QY	901	AADPKYDIPDSVIAFLRGLNPPGNGWPEPLTRALEGRSEKAPLTVPEEEQAHILDA	960
DB	901	AADPKYDIPDSVIAFLRGLNPPGNGWPEPLTRALEGRSEKAPLTVPEEEQAHILDA	960
QY	961	DDSKERNNSNLLLPKPTTEEFLEHRRRFGNTSALDDREFFYGLVREGRETLRLPDVRTP	1020
DB	961	DDSKERNNSNLLLPKPTTEEFLEHRRRFGNTSALDDREFFYGLVREGRETLRLPDVRTP	1020
QY	1021	LLVRLDAISEPDDKGMNVNANVNGQIRPMVRDRDSVESVTATAEKADSSNKGHVAAPFA	1080
DB	1021	LLVRLDAISEPDDKGMNVNANVNGQIRPMVRDRDSVESVTATAEKADSSNKGHVAAPFA	1080
QY	1081	GVWTVTVAGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS	1140
DB	1081	GVWTVTVAGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS	1140

RESULT 11
 ID AAB83180 standard; protein; 1139 AA.
 XX AAB83180;
 AC AAB83180;
 DT 09-JUL-2001 (first entry)
 DE Corynebacterium thermoaminogenes pc protein.
 XX Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
 KW thermotolerant; acbA; accBC; dtsR2; ptk; scrB; gluABCD; pdhA; pc;
 KW ppc; acn; icd; ipd; ohaA.
 XX Corynebacterium thermoaminogenes.
 OS WO200125447-A1.
 XX 12-APR-2001.
 XX 04-OCT-2000; 2000WO-JP006913.
 XX 04-OCT-1999; 99JP-00282716.
 PR 01-NOV-1999; 99JP-00311147.
 PR 21-APR-2000; 2000JP-00120687.
 XX (AJIN) AJINOMOTO CO INC.
 XX Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K;
 PI Kimura E, Osumi T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
 PI Sugimoto S;
 XX WPI: 2001-300170/31.
 DR N-PSDB; AAF87437.
 XX
 XX Proteins and their DNA useful for microbial production of L-amino acids.
 PS Claim 9; Page 132-135; 215pp; Japanese.
 XX
 CC The present sequence is provided in a specification relating to genes
 CC encoding thermophilic amino acid biosynthesis system enzymes of the
 CC thermotolerant bacterium Corynebacterium thermoaminogenes. The novel
 CC proteins retain at least 30% isocitrate lyase activity after heating at
 CC 50°C for 5 minutes. DNA fragments encoding the enzymes were isolated from
 CC a Corynebacterium thermoaminogenes chromosomal DNA plasmid library by
 CC PCR. The DNA may be used for developing strains of amino acid producing
 CC microorganisms
 XX
 SQ Sequence 1139 AA;
 Query Match 91.6%; Score 5302.5; DB 4; Length 1139;
 Best Local Similarity 91.0%; Pred. No. 0;
 Matches 1036; Conservative 43; Mismatches 58; Indels 1; Gaps 1;
 QY 3 THTSSSTLPAPFKKILVANRGEIARAFPALETGAATVAIYPRDRGSHRSPASEAVRIG 52
 DB 3 TTFPSSTLPAPFKKILVANRGEIARAFPALETGAATVAIYPRDRGSHRSPASEAVRIG 52
 QY 63 TEGSPVKAYLDIIEIIIGAAKKVADAIYPGYGLFSENAQLARECAENGITFIGTPEVLD 122
 DB 63 TEGSPVKAYLDIIEIIIGAAKKVADAIYPGYGLFSENAQLARECAENGITFIGTPEVLD 122
 QY 123 LTGDKSRVTAAKAGLPVLAETPSKNIDEIVKSAEGQTYPIFKAVAGGGGRGRFVA 182
 DB 123 LTGDKSRVTAAKAGLPVLAETPSKNIDEIVKSAEGQTYPIFKAVAGGGGRGRFVE 182
 QY 183 SPDELRLKATASREAAAFGDGAVYVERAVINPQHIIEVQILGDHTGEVVHLYERDCSLQ 242
 DB 183 KPDELRLAREASREAAAFGDGAVYVERAVINPQHIIEVQILGDHTGDVHLYERDCSLQ 242
 QY 243 RRHQKVEIAPAGHLDPELDRLCADAVKFCRIGYCGAGTVEFLVDEKGNHVFEMNPR 302

DB 243 RRHQKVEIAPAGHLDPELDRLCADAVKFCRIGYCGAGTVEFLVDEKGNHVFEMNPR 302
 QY 303 IQVEHTVTEEVTDLVKAQMRLAAGATLKELGHTODKIKTHGAALOCRIITTEPNNGCFR 362
 DB 303 IQVEHTVTEEVTSVDLVKAQMRHAAGATLKELGHTODKIKTHGAALOCRIITTEPNNGCFR 362
 QY 363 PDTGTTIAYRSPGGAGVRLDGAALGGEITAHFDSMLVMTCRGSDPETAVARAQRALAE 422
 DB 363 PDTGTTIAYRSPGGAGVRLDGAALGGEITAHFDSMLVMTCRGSDPETAVARAQRALAE 422
 QY 423 FTVSGVATNIGIFALLAREEDFTSKRTATGFIADHPHLLQAPPADDCGRILDYADVTV 482
 DB 423 FNVSGVATNIGIFALLAREEDFTSKRTATGFIADHPHLLQAPPADDCGRILDYADVTV 482
 QY 483 NKPGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDITFR 542
 DB 483 NKPGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDITFR 542
 QY 543 DAHQSLATRVRSFALTPAARAVAKLTPELLSVANGGATYDVAMRFLFEDPWRLDELRL 602
 DB 543 DAHQSLATRVRSFALTPAARAVAKLTPELLSVANGGATYDVAMRFLFEDPWRLDELRL 602
 QY 603 EAMPNVNIQMLLRGNTVGYTPYDPSVCRAFPVKEAASGVDFIFRFDALNDVSQMRPAID 662
 DB 603 EAMPNVNIQMLLRGNTVGYTPYDPSVCRAFPVKEAASGVDFIFRFDALNDVSQMRPAID 662
 QY 663 AVLENTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLRPA 722
 DB 663 AVLENTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLRPA 722
 QY 723 AVTKLVATLRREFDLPVHVHTHTAGGOLATYPAQAQADAVDAGASAPLSGTTSQPSLS 782
 DB 723 AVTKLVATLRREFDLPVHVHTHTAGGOLATYPAQAQADAVDAGASAPLSGTTSQPSLS 782
 QY 783 AIVAAFAHTRRDTGLSLAEVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN 842
 DB 783 AIVAAFAHTRRDTGLSLAEVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN 842
 QY 843 LRAQATLGLADRFELIEDNYAAVNEMLGRPTKVTSSKVVGDIALHLVAGVDPADFAA 902
 DB 843 LRAQATLGLADRFELIEDNYAAVNEMLGRPTKVTSSKVVGDIALHLVAGVDPADFAA 902
 QY 903 DPQKYDIPDSVIAFLRGLNPPGGWPEPLNTRALEGRSEKAPLTVPEEQAHLDADD 962
 DB 903 DPQKYDIPDSVIAFLRGLNPPGGWPEPLNTRALEGRSEKAPLTVPEEQAHLDADD 962
 QY 963 SKERNLSNRLLPKPTPEEFLEHRRRFGNTSALDDREFFVGLVGRGTLRLPDVTRTELL 1022
 DB 963 SKERNLSNRLLPKPTPEEFLEHRRRFGNTSALDDREFFVGLVGRGTLRLPDVTRTELL 1022
 QY 1023 VRLDAISEPDDKGMNVVNVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFAGV 1082
 DB 1023 VRLDAISEPDDKGMNVVNVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFAGV 1082
 QY 1083 VVTVVAEGDEVKAGDAVAITEAMKWEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
 DB 1083 VVTVVAEGDEVKAGDAVAITEAMKWEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
 RESULT 12
 ID ABU25961 standard; protein; 1141 AA.
 XX ABU25961;
 AC ABU25961;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by prokaryotic essential gene #11488.
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Corynebacterium diphtheriae.
 OS XX

P	N		WO200277183-A2.	
X	X			
P	D		03-OCT-2002.	
X	X			
P	F		21-MAR-2002; 2002WO-US009107.	
X	X			
P	R		21-MAR-2001; 2001US-00815242.	
X	X		06-SEP-2001; 2001US-00948993.	
P	R		25-OCT-2001; 2001US-00429235.	
X	X		08-FEB-2002; 2002US-00072851.	
P	R		06-MAR-2002; 2002US-0362699P.	
X	X		(ELIT-) ELITRA PHARM INC.	
P	A			
X	X		Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen XL, Zyskind JW;	
P	I		Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;	
X	X			
P	R		N-PSDB; ACA29831.	
X	X			
P	T		New antisense nucleic acids, useful for identifying proteins or screening	
X	X		PT for homologous nucleic acids required for cellular proliferation to	
P	T		isolate candidate molecules for rational drug discovery programs.	
X	X			
P	S		Claim 25; SEQ ID NO 53885; 1766pp; English.	
X	X			
P	C		The invention relates to an isolated nucleic acid comprising any one of	
X	C		the 6213 antisense sequences given in the specification where expression	
P	C		of the nucleic acid inhibits proliferation of a cell. Also included are:	
X	C		(1) a vector comprising a promoter operably linked to the nucleic acid	
P	C		encoding a polypeptide whose expression is inhibited by the antisense	
X	C		nucleic acid; (2) a host cell containing the vector; (3) an isolated	
P	C		polypeptide or its fragment whose expression is inhibited by the	
X	C		antisense nucleic acid; (4) an antibody capable of specifically binding	
P	C		the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
X	C		proliferation or the activity of a gene in an operon required for	
P	C		proliferation; (7) identifying a compound that influences the activity of	
X	C		the gene product or that has an activity against a biological pathway	
P	C		required for proliferation, or that inhibits cellular proliferation; (8)	
X	C		identifying a gene required for cellular proliferation or the biological	
P	C		pathway in which a proliferation-required gene or its gene product lies	
X	C		or a gene on which the test compound that inhibits proliferation of an	
P	C		organism acts; (9) manufacturing an antibiotic; (10) profiling a	
X	C		compound's activity; (11) a culture comprising strains in which the gene	
P	C		product is overexpressed or underexpressed; (12) determining the extent	
X	C		to which each of the strains is present in a culture or collection of	
P	C		strains; or (13) identifying the target of a compound that inhibits the	
X	C		proliferation of an organism. The antisense nucleic acids are useful for	
P	C		identifying proteins or screening for homologous nucleic acids required	
X	C		for cellular proliferation to isolate candidate molecules for rational	
P	C		drug discovery programs, or for screening homologous nucleic acids	
X	C		required for proliferation in cells other than S. aureus, S. typhimurium,	
P	C		K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of	
X	C		the target prokaryotic essential genes. Note: The sequence data for this	
P	C		patent did not form part of the printed specification, but was obtained	
X	C		in electronic format directly from WIPO at	
P	C		ftp.wipo.int/pub/published_pat_sequences	
X	X			
S	Q		Sequence 1141 AA;	
			Query Match 80.3%; Score 4647.5; DB 6; Length 1141;	
			Best Local Similarity 80.6%; Pred. No. 0;	
			Matches 917; Conservative 81; Mismatches 137; Indels 3; Gaps 2	
Q	Y	2	STHTASTLPAPKKILVANRGEIIVAFRAALETGAATVAIYPREDRGSHFRSFASEAVRI 61	
D	b	6	SRGTSTVNPNUSKILVANRGEIIVAFRAAFETGAATVAVPYPNEDRSNFRSFASEAVLI 65	
Q	Y	62	GTEGSPVKAYLIDIRIIIGAAKKVKADAIYPCGYFLSENAQLARECAENGITTTGGTPEVL 121	
D	b	66	GGGSANVAXYLDIRIIIFAAQTGDALYPCGYFLSENAQLARECAENGITTTGGTPEVL 125	
Q	Y	122	DUTGDKSRVATHAKKAGIPLVALESTPFSKNIDEIVKSAGQTTFIFVKAVAGGGGRMRF 181	

XX 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #19553.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Mycobacterium avium.
 XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 XX 25-OCT-2001; 2001US-0342923P.
 XX 08-FEB-2002; 2002US-00072851.
 XX 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 XX N-PSDB; ACA37896.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 XX for homologous nucleic acids required for cellular proliferation to
 XX PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 61950; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 XX the 6213 antisense sequences given in the specification where expression
 XX of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX (1) a vector comprising a promoter operably linked to the nucleic acid;
 XX encoding a polypeptide whose expression is inhibited by the antisense
 XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX polypeptide or its fragment whose expression is inhibited by the
 XX antisense nucleic acid; (4) an antibody capable of specifically binding
 XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX proliferation or the activity of a gene in an operon required for
 XX proliferation; (7) identifying a compound that influences the activity of
 XX the gene product or that has an activity against a biological pathway
 XX required for proliferation, or that inhibits cellular proliferation; (8)
 XX identifying a gene required for cellular proliferation or the biological
 XX pathway in which a proliferation-required gene or its gene product lies
 XX or a gene on which the test compound that inhibits proliferation of an
 XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX compound's activity; (11) a culture comprising strains in which the gene
 XX product is overexpressed or underexpressed; (12) determining the extent
 XX to which each of the strains is present in a culture or collection of
 XX strains; or (13) identifying the target of a compound that inhibits the
 XX proliferation of an organism. The antisense nucleic acids are useful for
 XX identifying proteins or screening for homologous nucleic acids required
 XX for cellular proliferation to isolate candidate molecules for rational
 XX drug discovery programs, or for screening homologous nucleic acids
 XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 XX the target prokaryotic essential genes. Note: The sequence data for this
 XX patent did not form part of the printed specification, but was obtained
 XX in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 1127 AA;

Query Match 65.5%; Score 3791.5; DB 6; Length 1127;
 Best Local Similarity 66.4%; Pred. No. 5,4e-271;
 Matches 750; Conservative 140; Mismatches 229; Indels 11; Gaps 7;

QY	14	KILVANRGEIAVRAFAALETGAATVAIYPPREDRGSHRSFASAVRIGTSGSPVKAYLD	73
DB	4	KVLVANRGEIAIRAPRAAYELEMATVAVPYVEDRNSVHRLKADDESYQIGEGHPVRAVLS	63
QY	74	IDEIIGAKKVKADAIYPCYGFLENQAARECHENGITTFICPTFEVLDLTDGDSRAVTA	133
DB	64	VDEIVGTALACADAIYPGYGFLENPDAAACAAAGITFVGPSAEVLELTDGDSRAIAA	123
QY	134	AKKAGLPVLABSTPSKNIDEIVKSAEGQTYPIFKAVAGGGGGRMRFVASFDELRKLAIE	193
DB	124	ARAAGLPVLABSTPSSTVSQELLSAAETWTFPLFKAVAGGGGGRMRRVTDGALAEALIE	183
QY	194	ASREAAAFGCGAVYVERAVINPOHIEVOILGDHTGEVHLYERDSCSLORHQKVEIAP	253
DB	184	ASREAAAFGCGAVYVERAVINPHIEVOILGDHTGEVHLYERDSCSLORHQKVEIAP	243
QY	254	AQHLDPELRDRICADAVKFCISIGYQAGTVEFVLDKGNHVFTEMNPRIOVEHTVTEEV	313
DB	244	APNLDPALRERI CADAVAFASISGYTCAGTVEFLLDERGNHVFTEMNPRIOVEHTVTEI	303
QY	314	TEVDLVKAMRLAAGATILKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGITAYRS	373
DB	304	TDVLVSAQLRIASQOTLEEIGLSQDSVIRGAALQCRITTEDPANGFRPDTGITAYRT	363
QY	374	PGGAGVRLDGAALQCGEITAHFDSMLVMTCRGSDFFETAVARAQALAEFTVSGVATNIG	433
DB	364	PGGAGIRLDGGTTLGAEISAHFDSMLIKLICRGDRDFTFAVERARRAAEFRIKGVSTNIP	423
QY	434	FLRALLREEDTSKRIATGFIADPHILLQAPPDDEQGRILDYLDVYLVNKHGVRPKDV	493
DB	424	FLQAVLDDPDPQAGRITTSIEQRPQLLTARSSADRGTKILNYLDVYLVNKHGVRPSAV	483
QY	494	AAPIKDLPNIKDLPL--PRGSRDLKOLGPAAPFARDLREQDQALAVTDTTFRDAHQSLLAT	551
DB	484	-YPDKLPDI-DLAASPAGSKQBLTBLGPERFAANLRSGGVGVDTTFRDAHQSLLAT	541
QY	552	RVRFAKPAABAVAKUTPELLSVEAMGATYDVAMFLFEDPMDRLDEREAMPNVNIQ	611
DB	542	RVRTSGLIKVAPYARTWPPQLLSVECMGATYDVALFLKEDPWERLTALREAMPNICLQ	601
QY	612	MLLGRNTVGTVPDSCVAFVKEASSGVDFRIFDALNDYSOMEPATDAVLETTAV	671
DB	602	MLLGRNTVGTVPETVTTFAVAETETGIDIFRIFDALNNVDSMPAPDAVRETGSAL	661
QY	672	AEVAMVSGDLSDPNEKLYTLDDYVLMKAEIIVKSGAHIAIKOMAGLRLFAATKLVATL	731
DB	662	AEVAMVSGDLSDPNEKLYTLDDYVLMKADQIVEAGAHVLAIKOMAGLRLFAATLVSA	721
QY	732	RREEDLPVHVHTHTAGQATYFAAAGADADVDGASAPLSGTTSPSLSAIVAAFAHT	791
DB	722	KSRFDLPVHVHTHTDPGGQATYVAAQAGADAVDGAAPLAGTTSQPALSSIVAAANT	781
QY	792	RRDTGLSLEAVSDLEPYEAVRGLYLPFESGTPQPTGRVYRHEIPGGQLSNLRAQAATLG	851
DB	782	BYDTGLSLPACVCDLEPYEALRVYAFESGLPAPTGRVYRHEIPGGQLSNLRAQAATLG	841
QY	852	LADRFELIEDNYAANVEMLRPKVTPTSSKVGDLALHLVAGVDPADPADPKYDIPD	911
DB	842	LGDFEFEDINAYAGADAILGHLVKVTPSSKVGDLALALVAGVGAQDFEADPSRYDIPD	901
QY	912	SVIAFLRGELGNPPGWPPEPLRTRALEGRSEK--APLTVPEBEEQAHLDADDSKERNRS	969
DB	902	SVIGFLRGELGDPGWPPEPLRTKALQGRPAKPEQLT--AEDEAALAPGAR-RQAA	957
QY	970	LNRLLPKPTPEEFLEHRRRNGNTSALDDREFFVLVEGRETLLRLPDVTRFLVRLDAIS	1029
DB	958	LNRLLPKPTPEEFLEHRESEYQDTSLSANQFFYGLRQGDSEHVEL-ERGVELLIGLEAIS	1016
QY	1030	EPDDKGNRVNVANVNGOIRPMVRDRSVESVTATAEKADSSNKGHVAAFPAGVVTVAE	1089
DB	1017	DPDESGKRTVMCLINGQLRPVWVDRSIAIDVPAEAKADPRANDPHIAAPFAGVVTVAEV	1076
QY	1090	GDEVKAGDAVAIIEMKNEATITIASVDGKIDRVVVVPAATKVEGDDLIVV	1139

Db 1077 GROVEAGOTIATEAMKMEAAVTSPKSGKVARIASRTAQVEGGDLLMVI 1126

RESULT 14

ABU36886
ID ABU36886 standard; protein; 1127 AA.
XX AC ABU36886;
XX

19-JUN-2003 (first entry)

Protein encoded by Prokaryotic essential gene #22413.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

Mycobacterium tuberculosis.

WO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0062699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI: 2003-029926/02.

N-PSDB; ACA40756.

New antisense nucleic acids, useful for identifying proteins or screening

for homologous nucleic acids required for cellular proliferation to

isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 64810; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation or for screening homologous nucleic acids for drug discovery programs, or for screening candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1127 AA;

Query Match 64.1%; Score 3712.5; DB 6; Length 1127;
Best Local Similarity 64.4%; Pred. No. 3.7e-265;
Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;

QY 12 FKKILVANRGEIATVAFRAALETGAATVAIYPREDRGSFHRSPASEAVRIGTEGSPVKAY 71
DB 2 FSKVLVANRGEIATVAFRAAYELGVGTVAIYPEDNSOHRUKADESYOIGDHPHVAY 61
QY 72 LDIETIIGAAKKVADAIYGYGFLSENAQLARECAENGITFIGTPEVLDITGDKSRV 131
DB 62 LSVDEIVATARRAGADAIYGYGFLSENPDLAAACAAAGISFVGPFAEVLGAGNKSRAI 121
QY 132 TAAKAGLPVLAESTPSKNIDEIVKSAEGQTPIFVKA VAGGGGRGMRVAVSPDELKLA 191
DB 122 AAAREAGLPVLMSSAPASVDELLSVAAGMPFLPVKAVAGGGGRGMRVAGDIALPEAI 181
QY 192 TEASREAAAFGGGAVYVERAVINPQHIEVQILGDHTGEVHLRYERDCSLQRHOKVVEI 251
DB 182 EAASREAESAFGDPVTYLEQAVINPRHIEVQILADNLGDVHLRYERDCSVQRHOKVIEL 241
QY 252 APAQHLDPELRDRICADAVKFCRSIGYCGAGTVFELVDEKGNHVFIEKMPRIQVSHVTE 311
DB 242 APAPHLDALRYKMCVDVAFARHIGYSCAGVFEFLDERGEYVFIEMPRVQVSHVTE 301
QY 312 EVTEVDLVRQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGITAY 371
DB 302 EITDVLVASQLRIAAGETLEQLGRQEDIAHPGAALQCRITTEDPANGFRPDTGRISAL 361
QY 372 RSPGGAGVLDGAAQOLGGEITAHEDSMVLMVTCRSDSETAVARAQALAEVTSVATN 431
DB 362 RTAGGAGVRLDGSNTLGAISPIFDSMLVKLCRCRDLPVAVSRARRAIEFRIQVSTN 421
QY 432 IGFLRALLREEDFTSKRIATGFIADPHLLOAPPADDEQGRILDYLDVYVKNKPGVVRPK 491
DB 422 IFFLOAVLDDPFRAGRVTTSIDERPQLLTARASADRGTKILNFLADVTVNNPYGSRPS 481
QY 492 DVNAAPIDKLPNIKDLPL-----PRGSRDLKQLGPAAPFARDLREQDALAVDTTFFDAHQ 547
DB 482 TI-YPDCKLPL---DLDLRAAPPAGSKQRLVKLGPEGFARWLRESAAVGVDTTFFDAHQ 537
QY 548 LLATRVRSFALKPAEAAVAKLTPELLSVEAMGATYDVAMREFLPDPDRDLDELREAMPN 607
DB 538 LLATRVRTSGLSRVAPYLARTMPQLLSVECGATYDVVALRFLKEDPHERLATLRAAMPN 597
QY 608 VNIOMLLRGNTVGYTPPDSVCRAFVKEAASSGVVDIFRIFDALNDVSMQPAIDAVLET 667
DB 598 ICLQMLLRGNTVGYTPPEIYTSFVQSEATATGIDIFRIFDALNNIESMPAIDAVRET 657
QY 668 NTAVAEVAMAYSGDLSDPNEKLYLDYLLKMAEIVKSGAHILAIKDMAGLLRPAAVTKI 727
DB 658 GSAIAEVAMCYTGDLTGCEQLYLDYLLKMAEQIVDAGAHVLAIKDMAGLLRPPAAQRL 717
QY 728 VTALRREFPLPVHVTHTDTAGQLATYFAAQAGADAVDASAPLSGTSQPSLSAIVAA 787
DB 718 VSALRSRFLPVHLNTHDTPGGQLASYAAAHAGADAVDGAAPLAGTTSQPALSSIVAA 777
QY 786 FATHRRDRTGLSLEAVSDLEPYWEAVRGLVLPPESTPGTGRVYRHEIPGGQLSNLRAQA 847
DB 778 AATHEYDTGLSLSAVCALEPYWEALRYAPPESELPGTGRVYRHEIPGGQLSNLROQA 837
QY 848 TALGLADRFELTEDNYAAVNEMLGRPTKVTPESSKVVGDALHLHVGAVDVPADFAADPOKY 907
DB 838 IALGLGDRPEEIEEAYAGADRVLGSLVKVTPPTSKVVGDALALALVAGVSADEFAFSDPARF 897
QY 908 DIPDSVIAFLRGELGNPPGGMPEPLRTALEGRSGKAPLTPVEPEQAHLDADDKERR 967
DB 898 GIPESVLGFLRGELGDPGGMPEPLRTAALAGRAAR-PTAQLAADDEIALLSSVGAK-RQ 955

QY 848 TALGLADREBELIEDNYAANEMIGRPTKVTPSSKVVGDIALHLVAGVDPADFAADPOKY 907
 Db 838 IALGLGDRFEIEEAYAGADRVLGRIVKVTPTSKVVGDIALALVAGVSADEFASDPARF 897
 QY 908 DIPDSVIAFLRGELGNPPGCGWPEPLRTRALEGRSEKGAPLTEVPPEEQAHLDADDKERR 967
 Db 898 GIPESVLGLRGELGDPGCGWPEPLETALAGRGAR-PTAQLAADDEIALSSVGAK-RQ 955
 QY 968 NSLNRLFFPKPTBEFLHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1027
 Db 956 ATLNRLFFSPTKFEFNEHEEAYGDTSQLSANOFFYGLRQGEHRVKL-ERGVELLIGLEA 1014
 QY 1028 ISEPDCKGMRNVVANVNGOIRPMRVDRDSVESVTATAEKADSSNKGHVAAPFAGVVTVV 1087
 Db 1015 ISEPDERGMRVWCILNGQLRPVLVDRDSIASAVPAAEKADRGNECHIAAPFAGVVTGV 1074
 QY 1088 ABGDEVKAGDAVAITIAMKOEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
 Db 1075 CVGERVGAGQTIATIAMKOEAPITAPVAGTVERVAVSQTAVQVEGGDLIVVVS 1127

Search completed: March 24, 2004, 22:42:35
 Job time : 69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:39:57 ; Search time 57 Seconds
(without alignments)
6310.364 Million cell updates/sec

Title: US-10-045-072-2

Perfect score: 5788

Sequence: 1 MSHTSTSTLPAFKILVANR.....RVVPAATKVEGDLIVVVS 1140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvrius:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5788	100.0	1140	16	O54587
2	5759	99.5	1140	2	Q8L2G4
3	5317.5	91.9	1168	16	Q8FRQ0
4	5305.5	91.7	1139	2	Q8RQJ2
5	3799	65.6	1124	16	Q9RK64
6	3783.5	65.4	1127	2	Q9F843
7	3712.5	64.1	1127	16	P95127
8	3712.5	64.1	1127	16	Q7TXJ1
9	3483	60.2	1124	2	O50450
10	3248	56.1	1131	16	Q83HF3
11	3241	56.0	1131	16	Q83H55
12	2552	44.1	1178	13	Q8JHF6
13	2540.5	43.9	1178	6	Q7YS28
14	2528	43.7	1148	16	Q81M9
15	2526	43.6	1148	16	Q81MT6
16	2535.5	43.6	1150	16	Q9K9M0

17	2525.5	43.6	1178	6	Q866R1
18	2524	43.6	1148	16	Q9KWU4
19	2519.5	43.5	1180	13	Q9DDT1
20	2510	43.4	1166	16	Q7UES1
21	2509.5	43.4	1192	3	Q9HES8
22	2505	43.3	1147	2	P94448
23	2495	43.1	1167	16	Q7WPS4
24	2493	43.1	1146	16	Q8Y846
25	2491.5	43.0	1144	16	Q97FR7
26	2490	43.0	1146	16	Q92CW1
27	2489	43.0	1150	16	Q8NX69
28	2488	43.0	1150	16	Q9UYU8
29	2487.5	43.0	1193	3	Q9318
30	2477.5	42.8	1158	16	Q8FYT0
31	2476	42.8	1167	16	Q7WST0
32	2473.5	42.7	1174	16	Q8UEX3
33	2470.5	42.7	1152	16	Q8FE27
34	2464.5	42.6	1158	16	Q8YU20
35	2464	42.6	1153	16	Q8CPM3
36	2464	42.6	1175	5	O17732
37	2463.5	42.6	1142	16	Q831P4
38	2462.5	42.5	1179	13	Q8AVN3
39	2443	42.2	1185	3	Q9UUE1
40	2441	42.2	1181	5	Q9XZ00
41	2437.5	42.1	1175	3	Q8X1T3
42	2435	42.1	1152	16	Q92L13
43	2435	42.1	1185	3	P78822
44	2435	42.1	1197	5	Q8MKW5
45	2419	41.8	1195	5	Q16921

ALIGNMENTS

RESULT 1

O54587 PRELIMINARY; PRT; 1140 AA.
AC O54587;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1).
GN PYC OR CGL0689.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
CC NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=21253;
RA Koffas M.A.G., Ramamoorthi R., Pine W.A., Sinskey A.J.,
RA Stephanopoulos G.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032;
RA Peters-Wendisch P.G., Kreutzer C., Kalinowski J., Patek M., Sahn H.,
RA Bismann B.J.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RL "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- COPACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; AF038548; AAB32588.1; -;
DR EMBL; Y09548; CAAY0739.1; -;
DR EMBL; AF005276; BAB98082.1; -;
DR HSSP; P24182; 1BNC.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003374; P:biotin binding; IEA.

DR	GO: 0016874; F.ligase activity; IEA.	
DR	GO: 0004736; F.pyruvate carboxylase activity; IEA.	
DR	GO: 0006094; F.gluconeogenesis; IEA.	
DR	GO: 0008152; P.metabolism; IEA.	
DR	InterPro: IPR001882; Biotin_BS.	
DR	InterPro: IPR005482; Biotin_carb_C.	
DR	InterPro: IPR000089; Biotin_lipoyl.	
DR	InterPro: IPR005479; Cpase_L_D2.	
DR	InterPro: IPR005481; Cpase_L_N.	
DR	InterPro: IPR000891; HMG_L-like.	
DR	InterPro: IPR003379; PYC_OADA.	
DR	InterPro: IPR005930; Pyruv_carbox.	
DR	Pfam: PF02785; Biotin_carb_C; 1.	
DR	Pfam: PF00364; biotin_lipoyl; 1.	
DR	Pfam: PF00289; Cpase_L_chain; 1.	
DR	Pfam: PF02786; Cpase_L_D2; 1.	
DR	Pfam: PF00682; HMG_L-like; 1.	
DR	Pfam: PF02436; PYC_OADA; 1.	
DR	TIGRfam: TIGR01235; pyruv_carbox; 1.	
DR	PROSITE: PS00188; BIOTIN; 1.	
DR	PROSITE: PS00867; CPASE_2; 1.	
KW	Biotin; Ligase; Pyruvate; Complete; proteome.	
Q	SEQUENCE 1140 AA; 123102 MW; 2A6D4B4ED2FEB531 CRC64;	
Query Match 100.0%; Score 5788; DB 16; Length 1140;		
Best Local Similarity 100.0%; Pred. No. 1.4e-299;		
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MSTHTSSSTLPAPFKKILVNRGIAVRAFAALETGAATVAIYPRDRGSPHRSFASAVR 60	
DB	1 MSTHTSSSTLPAPFKKILVNRGIAVRAFAALETGAATVAIYPRDRGSPHRSFASAVR 60	
QY	61 IGTEGSPVKAYLDIDBIIGAANKVADAIYPGYGLSNAQLARECAENGTFITGPTPEV 120	
DB	61 IGTEGSPVKAYLDIDBIIGAANKVADAIYPGYGLSNAQLARECAENGTFITGPTPEV 120	
QY	121 LOLTGDKSAVTAANKAGLPVLAESTPSKNIDEIVKSAEGOTYPIFKVAGGGRGNRF 180	
DB	121 LOLTGDKSAVTAANKAGLPVLAESTPSKNIDEIVKSAEGOTYPIFKVAGGGRGNRF 180	
QY	181 VASPELRLKLAETASREAAFGDGAIVYERAVINPQIHVEQILGDHTGEVHHLYERDCS 240	
DB	181 VASPELRLKLAETASREAAFGDGAIVYERAVINPQIHVEQILGDHTGEVHHLYERDCS 240	
QY	241 LQRRHQKVEIAPQHLPELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFEMN 300	
DB	241 LQRRHQKVEIAPQHLPELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFEMN 300	
QY	301 PRIQVEHTVEVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360	
DB	301 PRIQVEHTVEVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360	
QY	361 FRPDTGTTIAYRSPGAGVRLDGAALGGEITAHFDSMLVMTCRGSDFTAVARAQAL 420	
DB	361 FRPDTGTTIAYRSPGAGVRLDGAALGGEITAHFDSMLVMTCRGSDFTAVARAQAL 420	
QY	421 ASFTVSGVATNIGFRLALREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDLADY 480	
DB	421 ASFTVSGVATNIGFRLALREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDLADY 480	
QY	481 TYNKCHGVTPKVAAPIDKIPNIDLPGRSPDLKOLGPAAPARDIREQDALAVDTT 540	
DB	481 TYNKCHGVTPKVAAPIDKIPNIDLPGRSPDLKOLGPAAPARDIREQDALAVDTT 540	
QY	541 FRDAHQSLLATVRSPALKFAAEAVAKLTPELLSVGAGGATYDVAMRFLFEDPWRLDE 600	
DB	541 FRDAHQSLLATVRSPALKFAAEAVAKLTPELLSVGAGGATYDVAMRFLFEDPWRLDE 600	
QY	601 LREAMPNNIOMLLRGENTVGTTPYDPSVCRAFKVEAASSGVDIPRIFDALNDVSQMRPA 660	
DB	601 LREAMPNNIOMLLRGENTVGTTPYDPSVCRAFKVEAASSGVDIPRIFDALNDVSQMRPA 660	
QY	661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKVAEEIVKSGAHILAIDKMAGLLR 720	

DB	661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKVAEEIVKSGAHILAIDKMAGLLR 720	
QY	721 PAAVTKLVTLARRFDPVHVHTHTAGGLATYFAAAQAGADAVDGSAPLSGTTSPS 780	
DB	721 PAAVTKLVTLARRFDPVHVHTHTAGGLATYFAAAQAGADAVDGSAPLSGTTSPS 780	
QY	781 LSAIVAFAFHTRRDTGLSLEAVSDLEPYWEAVRGYLPESGTPGPTGVYHEIPGGOL 840	
DB	781 LSAIVAFAFHTRRDTGLSLEAVSDLEPYWEAVRGYLPESGTPGPTGVYHEIPGGOL 840	
QY	841 SNLRAQATGALADREFELIEDNYAAVNEMLGPTKVTPTSSKVVGDALHLVAGVDPADF 900	
DB	841 SNLRAQATGALADREFELIEDNYAAVNEMLGPTKVTPTSSKVVGDALHLVAGVDPADF 900	
QY	901 AADPQYDIPDSVIAFLRGELGNPPGWPPEPLRTALRGSRSEKAPLTPVPEEQAHLDA 960	
DB	901 AADPQYDIPDSVIAFLRGELGNPPGWPPEPLRTALRGSRSEKAPLTPVPEEQAHLDA 960	
QY	961 DSKERNLSNRLLPFKPTTEFLHRRRFGNTSALDDREFFYGLVEGRETLLRLFDVTRP 1020	
DB	961 DSKERNLSNRLLPFKPTTEFLHRRRFGNTSALDDREFFYGLVEGRETLLRLFDVTRP 1020	
QY	1021 LLVRLDAISEPDDKGNRVVAVNNGQIRPMRVDRSRVSVTATAEKADSSNKGHVAAPFA 1080	
DB	1021 LLVRLDAISEPDDKGNRVVAVNNGQIRPMRVDRSRVSVTATAEKADSSNKGHVAAPFA 1080	
QY	1081 GVVTVTVAGDEVKAGDAVAIIEMKWEATITASVDGKIDRVVWPAATKVEGDDLIVVVS 1140	
DB	1081 GVVTVTVAGDEVKAGDAVAIIEMKWEATITASVDGKIDRVVWPAATKVEGDDLIVVVS 1140	
RESULT 2		
Q8L2G4	PRELIMINARY; PRT; 1140 AA.	
ID	Q8L2G4;	
AC	Q8L2G4;	
DT	01-OCT-2002 (T-EMBLrel. 22, Created)	
DT	01-OCT-2002 (T-EMBLrel. 22, Last sequence update)	
DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)	
DE	Pyruvate carboxylase.	
GN	PyC.	
OS	Corynebacterium crenatum.	
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.	
OX	NCBI_TaxID:168810;	
RN	[1]_TaxID:168810;	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CD945;	
RA	Wang J., Ding J., Liu Y.;	
RT	"Cloning and Expression of Pyruvate Carboxylase Gene in	
RT	Corynebacterium crenatum CD945."	
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AF503915; AM27458.1; -	
DR	GO: 0005737; C.cycloplasm; IEA.	
DR	GO: 0005524; F.ATP binding; IEA.	
DR	GO: 0009374; F.biotin binding; IEA.	
DR	GO: 0016874; F.ligase activity; IEA.	
DR	GO: 0004736; F.pyruvate carboxylase activity; IEA.	
DR	GO: 0006094; F.gluconeogenesis; IEA.	
DR	GO: 0008152; P.metabolism; IEA.	
DR	InterPro: IPR001882; Biotin_BS.	
DR	InterPro: IPR005482; Biotin_carb_C.	
DR	InterPro: IPR000089; Biotin_lipoyl.	
DR	InterPro: IPR005479; Cpase_L_D2.	
DR	InterPro: IPR005481; Cpase_L_N.	
DR	InterPro: IPR000891; HMG_L-like.	
DR	InterPro: IPR003379; PYC_OADA.	
DR	InterPro: IPR005930; Pyruv_carbox.	
DR	Pfam: PF02785; Biotin_carb_C; 1.	
DR	Pfam: PF00364; biotin_lipoyl; 1.	
DR	Pfam: PF00289; Cpase_L_chain; 1.	
DR	Pfam: PF02786; Cpase_L_D2; 1.	
DR	Pfam: PF00682; HMG_L-like; 1.	

DR Pfam: PF02436; PYC OADA: 1.
DR TIGRFAMS; TIGR01235; Pyruvate carboxylase; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPASE_2; 1.
SQ SEQUENCE 1140 AA; 123126 MW; 123126 MW; PFA90BB7644C910E CRC64;

Query Match 99.5%; Score 5759; DB 2; Length 1140;
Best Local Similarity 99.3%; Pred. No. 5.1e-298;
Matches 1132; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTHTSTLPAPFKILVANRGEIAPRAALETGAATVAIYPRDRGSHFSPASEAVR 60
DB 1 MSTHTSTLPAPFKILVANRGEIAPRAALETGAATVAIYPRDRGSHFSPASEAVR 60

QY 61 IGTGSPVKAYLIDIIIGAAKKVADAIYPGVFLSENAQLARECAENGITFIPTPEV 120
DB 61 IGTGSPVKAYLIDIIIGAAKKVADAIYPGVFLSENAQLARECAENGITFIPTPEV 120

QY 121 LDLTGKSAVTAAKKAGLPVLAESTPSKNIDIVKSAEGQYPIFVKAVAGGGRGMRP 180
DB 121 LDLTGKSAVTAAKKAGLPVLAESTPSKNIDIVKSAEGQYPIFVKAVAGGGRGMRP 180

QY 181 VASPELRLKLTASREAAAFDGAIVYVERAVINPQHIETVILGDHGTGEVHLVERCS 240
DB 181 VASPELRLKLTASREAAAFDGAIVYVERAVINPQHIETVILGDHGTGEVHLVERCS 240

QY 241 LQRHOKVVEIAPAOHLDPEDLDRI CADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEVN 300
DB 241 LQRHOKVVEIAPAOHLDPEDLDRI CADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEVN 300

QY 301 PRQVHTVTEETVDLKAQMLAAGATLKEGLTQDKIKTHGAALOCRTTDPNNG 360
DB 301 PRQVHTVTEETVDLKAQMLAAGATLKEGLTQDKIKTHGAALOCRTTDPNNG 360

QY 361 FREDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDPETAVARAQRAL 420
DB 361 FREDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDPETAVARAQRAL 420

QY 421 AETVSGVATNIGFLRALREEDFTSKRTATGFIADPHLLQAPPADDEQGRILDYADV 480
DB 421 AETVSGVATNIGFLRALREEDFTSKRTATGFIADPHLLQAPPADDEQGRILDYADV 480

QY 481 TVNKPGRVKPDVAAPIDKLPNKLPLRGRSDRLKQLGPAFAFARLREQALAVTDIT 540
DB 481 TVNKPGRVKPDVAAPIDKLPNKLPLRGRSDRLKQLGPAFAFARLREQALAVTDIT 540

QY 541 FRDAHQSLATRVRSFALPAAVAKLTPELLISVEANGATYDVAMRFLFDPDWRLDE 600
DB 541 FRDAHQSLATRVRSFALPAAVAKLTPELLISVEANGATYDVAMRFLFDPDWRLDE 600

QY 601 LREAMPNVIOMLGRNTVGYTPYDPSVCRAFPVKEAASGVVDIPRI PDALNDVSMRPA 660
DB 601 LREAMPNVIOMLGRNTVGYTPYDPSVCRAFPVKEAASGVVDIPRI PDALNDVSMRPA 660

QY 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTDYLLKMAEIVKSGAHLAIKDWAGLLR 720
DB 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTDYLLKMAEIVKSGAHLAIKDWAGLLR 720

QY 721 PAAVKLVTLALREEDFLPVHVTHTDAGGLATYFAAAGADAVDNGASAPLSGTTSPS 780
DB 721 PAAVKLVTLALREEDFLPVHVTHTDAGGLATYFAAAGADAVDNGASAPLSGTTSPS 780

QY 781 LSAIVAFAHTRDRTGLSLEAVSDLEPYNEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
DB 781 LSAIVAFAHTRDRTGLSLEAVSDLEPYNEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840

QY 841 SNLRAQATLGLADRFELIEDNYAAVNEMLGRPTKVTTPSKVVGVDLALHLVAGVDPADF 900
DB 841 SNLRAQATLGLADRFELIEDNYAAVNEMLGRPTKVTTPSKVVGVDLALHLVAGVDPADF 900

QY 901 AADPQKYDIPDSVIAFLRGELGNPPGWPPELPTALRSEGRSEKAPLTVPPEEQAHLLDA 960
DB 901 AADPQKYDIPDSVIAFLRGELGNPPGWPPELPTALRSEGRSEKAPLTVPPEEQAHLLDA 960

QY 961 DDSKERNLSNRLILFPKPTBEFLEHRRRFRONTSALDDREFFYGLVGEETLIRLPDVRTP 1020
DB 961 DDSKERNLSNRLILFPKPTBEFLEHRRRFRONTSALDDREFFYGLVGEETLIRLPDVRTP 1020

QY 1021 LLVRLDAISEPDDKGMENVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
DB 1021 LLVRLDAISEPDDKGMENVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

QY 1081 GWTVTVAEGDEVKAGDAVAIIEMKMEATITASVKGKIDRVVPAATKVEGGDLIVWS 1140
DB 1081 GWTVTVAEGDEVKAGDAVAIIEMKMEATITASVKGKIDRVVPAATKVEGGDLIVWS 1140

RESULT 3
Q8PRQO PRELIMINARY; PRT; 1168 AA.

ID Q8PRQO
AC Q8PRQO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pyruvate carboxylase (BC 6.4.1.1).
GN PYC OR C80703.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikao K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005216; BAC17519.1;
DR GO; GO:0005737; C:cyttoplasm; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; PYC_OADA.
DR InterPro; IPR005930; Pyruv carbox.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRFAMS; TIGR01235; pyruv carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPASE_2; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 1168 AA; 126245 MW; ASD5A4DD4D4D285F8 CRC64;

Query Match 91.9%; Score 5317.5; DB 16; Length 1168;
Best Local Similarity 91.8%; Pred. No. 1.7e-274;
Matches 1038; Conservative 43; Mismatches 56; Indels 1; Gaps 1;

QY 3 THTSTLPAPFKILVANRGEIAPRAALETGAATVAIYPRDRGSHFSPASEAVRIG 62
DB 32 TTTTSTLPAPFKILVANRGEIAPRAALETGAATVAIYPRDRGSHFSPASEAVRIG 91
QY 63 TEGSPVKAYLIDIIIGAAKKVADAIYPGVFLSENAQLARECAENGITFIPTPEVLD 122

Db	92	TEGSPVKAYLDIDEINAAKVKADAVPGYGLSENAQLARECAENGITFIPTPEVL	151
Qy	123	LTGDKSRAVTAAKKAGLPVLAESTPSKNIDELKSAEGQTYPIFKVAGGGGGMFEVA	182
Db	152	LTGDKSKAVSAKKAGLPVLAESTSTDEIDELKSAEGQTYPIFKVAGGGGGMFEVA	211
Qy	183	SPDELKXLAETASREAAAFGPGAYVVERAVINPQHIEVQILGDHTGEVHLYERDCSLQ	242
Db	212	KPEDLRELAREASREAAAFGPGSVYVERAVIKPQHIEVQILGDHTGDVHLYERDCSLQ	271
Qy	243	RHOKVVELAPAHQLDPELDRI CADAVKFCGSGYQAGATVEFLVDEKGNHVFIEKNR	302
Db	272	RHOKVVELAPAHQLDPELDRI CADAVKFCGSGYQAGATVEFLVDEKGNHVFIEKNR	331
Qy	303	IOVEHTVTEEVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALOCRIITDPNNGFR	362
Db	332	IOVEHTVTEEVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALOCRIITDPNNGFR	391
Qy	363	PTGVTITAVRPGGAGVLDGAQOLGGEITAHFDSMLVKMTCRSGDDETAVARQALAE	422
Db	392	PTGVTITAVRPGGAGVLDGAQOLGGEITAHFDSMLVKMTCRSGDDETAVARQALAE	451
Qy	423	FTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYLDVTV	482
Db	452	FNVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILEYLDVTV	511
Qy	483	NKPHGVTRPKDVAAPDKLPNIDLPGRSDRLKQLGPAFARDLREODALATVDTTFR	542
Db	512	NKPHGERP-ETARPTEKLEPEVNIPLPRGSRDLKQLGPEFARDLREODALATVDTTFR	570
Qy	543	DAHQSLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFEDPWRDLDEL	602
Db	571	DAHQSLATRVRSFALTPAARAVAKLTPELLSVEAWGGATYDVAMRFEDPWRDLDEL	630
Qy	603	EMPNVNQMLLRGNTVGYTPYPSVCRFAFKBAASSGVDFRIFDALNDVDSQWRPAID	662
Db	631	EMPNVNQMLLRGNTVGYTPYPSVCRFAFKBAASSGVDFRIFDALNDVDSQWRPAID	690
Qy	663	AVLENTAVAEVAMVSGDLSNPENKLYTLDYLLKMAEIIKSGAHILAIKDMAGLRLPA	722
Db	691	AVLENTAVAEVAMVSGDLSNPENKLYTLDYLLKMAEIIKSGAHILAIKDMAGLRLPA	750
Qy	723	AVTKLVTLARRFDLPVHVHTHTAGGQATYFAAAQAGADAVGASAPLSGTTQSQSL	782
Db	751	AAATKLVTLARRFDLPVHVHTHTAGGQATYFAAAQAGADAVGASAPLSGTTQSQSL	810
Qy	783	ATVAAFAHTRRTDGLSLEAVSLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIFGGQLSN	842
Db	811	ALVAAFAHTRRTDGLNLOAVSLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIFGGQLSN	870
Qy	843	LRAQATAGLADRFELIEDYAAVNMELGRPTKVTTPSSKVYGDALHLVAGVDPADFAA	902
Db	871	LRAQAVAGLADRFELIEDYAAVNMELGRPTKVTTPSSKVYGDALHLVAGVSPEDFAA	930
Qy	903	DPQKVDIPDSVIAFLRGELGNPPGQWPEPLRTRALEGRSEKAPLTPVPEEQAHLDAD	962
Db	931	DPQKVDIPDSVIAFLRGELGNPPGQWPEPLRTRALEGRSQKAPLAEIPAEQAHLDSD	990
Qy	963	SKERNSNRLLPFKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPL	1022
Db	991	SAERRGTLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLKEGREELISLTGVSTPM	1050
Qy	1023	VRLDAISEPDDKGMNVNVTNGQIRPMRVRDRSVESVTATAEKADSNKGVHVAAPAG	1082
Db	1051	VRLDAVSPDDKGMNVNVTNGQIRPIKVRDRSVESVTATAEKADATNKGHVHVAAPAG	1110
Qy	1083	VTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS	1140
Db	1111	VTVTVAEGDEIKAGDAVAIIEMKMEATITAPVDGVIDRVVVVPAATKVEGGDLIVVVS	1168

RESULT 4
QBRQL2

ID	QBRQL2	PRELIMINARY;	PRT; 1139 AA.
AC	QBRQL2;		
DT	01-JUN-2002 (T-EMBLrel. 21, Created)		
DT	01-JUN-2002 (T-EMBLrel. 21, Last sequence update)		
DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)		
DE	Pyruvate carboxylase.		
GN	PYC.		
OS	Corynebacterium efficiens.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.		
OX	NCBI_TaxID=152794;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Akiyoshi N., Nonaka G., Kimura E., Kawahara Y., Sugimoto S.;		
RT	"Corynebacterium efficiens pyruvate carboxylase (pyc) gene, complete		
RT	CDS."		
RL	Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.		
CC	-1- COFACTOR: BIOTIN (BY SIMILARITY).		
DR	EMBL; AB083299; BAB8903.1; -		
DR	GO; GO:0005737; C:cytoplasm; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0009374; F:biotin binding; IEA.		
DR	GO; GO:0016874; F:ligase activity; IEA.		
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.		
DR	GO; GO:0006094; P:gluconeogenesis; IEA.		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	InterPro; IPR001882; Biotin_BS.		
DR	InterPro; IPR005482; Biotin_carb_C.		
DR	InterPro; IPR000089; Biotin_lipoyl.		
DR	InterPro; IPR005479; CPase_L_D2.		
DR	InterPro; IPR005481; CPase_L_N.		
DR	InterPro; IPR000891; HMGL-Like.		
DR	InterPro; IPR003379; PYC_OADA.		
DR	InterPro; IPR005930; Pyruv_carbox.		
DR	Pfam; PF02785; Biotin_carb_C; 1.		
DR	Pfam; PF00364; biotin_lipoyl; 1.		
DR	Pfam; PF00289; CPase_L_chain; 1.		
DR	Pfam; PF02786; CPase_L_D2; 1.		
DR	Pfam; PF00682; HMGL-Like; 1.		
DR	Pfam; PF02436; PYC_OADA; 1.		
DR	TIGRfams; TIGR01235; pyruv_carbox; 1.		
DR	PROSITE; PS00188; BIOTIN; 1.		
DR	PROSITE; PS00867; CPASE_2; 1.		
KW	Biotin.		
SK	SEQUENCE 1139 AA; 123068 MW; BA7023134519FAAA CRC64;		

Query Match 91.7%; Score 5305.5; DB 2; Length 1139;
Best Local Similarity 91.0%; Pred No. 7,1e-274;
Matches 1037; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

Qy	1	MSHTSSTLPAPKKILVANRGEIAPRAALETGAATVAIYPREDRGSFHRSFASEAVR	60
Db	1	MVTTTSTLPAPKKILVANRGEIAPRAALETGAATVAIYPREDRGSFHRSFASEAVR	60
Qy	61	IGTEGSPVKAYLDIDEIGAANKKADATYPCYGLSENAQLARECAENGITFIPTPEV	120
Db	61	IGTEGSPVKAYLDIDEIINAANKKADATYPCYGLSENAQLARECAENGITFIPTPEV	120
Qy	121	LDLTGKSRVATAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFKVAGGGGGMRF	180
Db	121	LDLTGKSKAVSAAKKAGLPVLAESTPTDIDEIVKSAEGQTYPIFKVAGGGGGMRF	180
Qy	181	VASPDRLKLAETASREAAAFGPGSVYVERAVINPQHIEVQILGDHTGEVHLYERDCS	240
Db	181	VEKPEDRLAREASREAAAFGPGSVYVERAVINPQHIEVQILGDHTGDVHLYERDCS	240
Qy	241	LQRRHQVVEIAPAHQLDPELDRI CADAVKFCRSGYQAGTVEFLVDEKGNHVFIEKN	300
Db	241	LQRRHQVVEIAPAHQLDPELDRI CADAVKFCRSGYQAGTVEFLVDEAGNHVFIEKN	300
Qy	301	PRIOEHTVTEEVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALOCRIITDPNNG	360
Db	301	PRIOEHTVTEEVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALOCRIITDPNNG	360

```

QY 361 FRPDTGTTAYRSPGAGVRLDGAALGGGTEITAHFDSMLVKMTCRGSDPETAVARAQAL 420
DB 361 FRPDTGTTAYRSPGAGVRLDGAALGGGTEITAHFDSMLVKMTCRGSDPETAVARAQAL 420
QY 421 AFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHPLLQAPPADDEQGRILYADV 480
DB 421 AFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHPLLQAPPADDEQGRILYADV 480
QY 481 TVNKPVGVPKVAAPIDKLPNIDKPLPGRSGDRKLQGPAAFAFARDLREQDALAVTDTT 540
DB 481 TVNKPVGVPKVAAPIDKLPNIDKPLPGRSGDRKLQGPAAFAFARDLREQDALAVTDTT 540
QY 541 FRDAHQSLLATRVRSFALKFAAFAVAKLTPELLSSVEAMGGATYDVAMRFLFEDPDWRLDE 600
DB 541 FRDAHQSLLATRVRSFALKFAAFAVAKLTPELLSSVEAMGGATYDVAMRFLFEDPDWRLDE 600
QY 601 LREAMPNVNIOMLLRGNTVGYTPYDPSVCRAPFVKEAASGVVDIIFRIFDALNDVQMPPA 660
DB 601 LREAMPNVNIOMLLRGNTVGYTPYDPSVCRAPFVKEAASGVVDIIFRIFDALNDVQMPPA 660
QY 661 IDAVLETTNATAVAEVAAYSGDLSNPKYLYDYLYLMAEEIVKSAHILAIKDMAGLLR 720
DB 661 IDAVLETTNATAVAEVAAYSGDLSNPKYLYDYLYLMAEEIVKSAHILAIKDMAGLLR 720
QY 721 PAAVTKLVTRREFDLPVHVHTHTAGGOLATYFAAAQAGADAVDGASAPLSGTTSPS 780
DB 721 PAAVTKLVTRREFDLPVHVHTHTAGGOLATYFAAAQAGADAVDGASAPLSGTTSPS 780
QY 781 LSAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGVYRHEIPGGOL 840
DB 781 LSAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGVYRHEIPGGOL 840
QY 841 SNLRAQATLGLADRELIEDNVAANVEMGLRPTKTPSSKVVGDIALHLVAGVDPADP 900
DB 841 SNLRAQATLGLADRELIEDNVAANVEMGLRPTKTPSSKVVGDIALHLVAGVDPADP 900
QY 901 AADPQKYDIPDSVIAFLRGLNPGGWPPELRLTRALEGRSGKAPLTPVPEEECAHLDA 960
DB 901 AADPQKYDIPDSVIAFLRGLNPGGWPPELRLTRALEGRSGKAPLTPVPEEECAHLDA 960
QY 961 DDSKERNLNRLLPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETILRLDVRTP 1020
DB 961 DDSKERNLNRLLPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETILRLDVRTP 1020
QY 1021 LVLRLDAISPDPKGMNVVANNVNGQIRPKVDRSVESVTATAEKADSNKHVAAPPA 1080
DB 1021 LVLRLDAISPDPKGMNVVANNVNGQIRPKVDRSVESVTATAEKADSNKHVAAPPA 1080
QY 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKVEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
DB 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKVEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
QY 1139 GVVTVTVAEGDEVKAGDAVAIIEAMKVEATITAPVDGVIDRVVVPAAATKVEGGDLIVVVS 1139
DB 1139 GVVTVTVAEGDEVKAGDAVAIIEAMKVEATITAPVDGVIDRVVVPAAATKVEGGDLIVVVS 1139

RESULT 5
Q9RK64 PRELIMINARY; PRT; 1124 AA.
ID Q9RK64
AC Q9RK64;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Pyruvate carboxylase.
GN SC00346 OR SCF11.26C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrall B.G., Rajandream M.A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kleser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RA "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RN Mol. Microbiol. 21:77-96(1996).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.D., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939105; CAB59603.1; -.
DR HSSP; P24182; 1BNC
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005482; Biotin carb C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR005479; Cphase_L_D2.
DR InterPro; IPR005481; Cphase_L_N.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; Pyc OADA.
DR InterPro; IPR005930; Pvrut carbox.
DR Pfam; PF02785; Biotin carb C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; CPhase_L_chain; 1.
DR Pfam; PF02786; CPhase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; Pyc OADA; 1.
DR TIGRfams; TIGR01235; pvrut carbox; 1.
DR PROSITE; PS00867; CPhase_2; 1.
KW Complete proteome.
SQ SEQUENCE 1124 AA; 121071 MW; 1FEF4C4643954D31 CRC64;

Query Match 65.6%; Score 3799; DB 16; Length 1124;
Best Local Similarity 65.2%; Pred.No.1.ie-193;
Matches 736; Conservative 154; Mismatches 228; Indels 10; Gaps 6;

QY 12 FKTLVANRGEIATRAFRAGYELGARTVAVPHEDRNLRLKADAEYELGEQHPVAY 61
DB 12 FKTLVANRGEIATRAFRAGYELGARTVAVPHEDRNLRLKADAEYELGEQHPVAY 61
QY 72 LDDEIIIAAKKVKADAIYFGYGLFSLSENAQLACEAENGITFTGPTPEVLDLTGDKSRV 131
DB 72 LDDEIIIAAKKVKADAIYFGYGLFSLSENAQLACEAENGITFTGPTPEVLDLTGDKSRV 131
QY 62 LSVEIVRAARRAGADAVYFGYGLFSLSENPALARACEAGITFVGPSARILELTGNKARV 121
DB 62 LSVEIVRAARRAGADAVYFGYGLFSLSENPALARACEAGITFVGPSARILELTGNKARV 121
QY 132 TAARKAGLPVLAESTPSKNIDEIVKSAEGOTYPTFVKAVAGGGGRGRFVASPDELKLA 191
DB 132 TAARKAGLPVLAESTPSKNIDEIVKSAEGOTYPTFVKAVAGGGGRGRFVASPDELKLA 191
QY 122 AAAREAGVPVLGGAPSTDDVELVRAADDVGFVFKAVAGGGGRGRFVASPDELKLA 181
DB 122 AAAREAGVPVLGGAPSTDDVELVRAADDVGFVFKAVAGGGGRGRFVASPDELKLA 181
QY 192 TEASREAAAFDGDVAVYVERAVINPQHIIEVQILGDHTGEVYVHLVYERDCSLQRHKKYVEI 251
DB 192 TEASREAAAFDGDVAVYVERAVINPQHIIEVQILGDHTGEVYVHLVYERDCSLQRHKKYVEI 251

```


Db 182 BAAGREASAFSGSTVFEKAVVEPRHIEVQIADGEGDVHIFERDCSVQRHQRVIEL 241
Qy 252 APAQHLPDLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFTEMNPRIQVEHTVTE 311
Db 242 APANLDPALRERICADAVNFARIGYRNAGTVEFLVDRDGNHVFTEMNPRIQVEHTVTE 301
Qy 312 EVTEVDLAVKQMLAAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGFRPDGTITAY 371
Db 302 EVTDVULVQSLRTAAGCTLADJLAGENTLKGALQCRITTEDPANGFRPDGTGQISAY 361
Qy 372 RSPGAGVRLD-GAAQJGGGRTAHFDSMLVKMTCRGSDFETAVARACRALAEFTVSGVAT 430
Db 362 RSPGSGIRLDGGTTHAGTEISAHFDSMLVKLSRCGRDFTTAVNARRAAVEPRIRGVAT 421
Qy 431 NIGFLRALLREEDFTSKRIATGPTADPHILLQAPPADDEQGRILDVLYADVTKNPKGVPR 490
Db 422 NIPFLQAVLDDPDPQAGRTVTFIEQRPHTLTARHSADRGTULTYLVADVTKNPKGPR 481
Qy 491 KDVAAPDKIPNIXDLPLPRGSRDLKQLPAPAFARDLRQDALAVTDFTFRDAHQSLIA 550
Db 482 -ELVDPLTKLPTAGSAGPPAGSRQLLAELGPEGFARLRESSTIGTVDFTFRDAHQSLIA 540
Qy 551 TRVTSFALKPAEAVAKLTPELLSVBANGGATYDVANRFLFEDBWDLDELREAMPVNI 610
Db 541 TRVTKMLAVPVVARTLPLLSELCWGGATYDVALLFLAEDPWERLALREAVENLCL 600
Qy 611 QMLRGRNTVGYTPYDPSVCRFAVKEAASSGVDIFRIFDALNDVQMRPAIDAVLETNTA 670
Db 601 QMLRGRNTVGYTPYFTEVTDFAVQEAATGIDIFRIFDALNDVQMRPAIEAVRGTGSA 660
Qy 671 VAEVAMYSGLSPNEKLYTLDYLLKMAEIVKSGAHILAIDKDMAGLLRPAATKLVTA 730
Db 661 VAEVALCYTADLSPSELYTLDYLLKMAEIVNAGARVLAVKDMAGLLRPAATLVSA 720
Qy 731 LRREFDLVHVHTHTAGQGLATYFAAAQACADAVDGASAPLSTQTSQPSLSAIVAAFAH 790
Db 721 LRREFDLVHLHTHTTGGQATYLAALQAGADAVDGNVASMAGTTQPSLSAIVAAFDH 780
Qy 791 TRRTDGLSLEAVDLEPYWEAVRGLYLPFESGTPGTRVVRHPIPGQJLSNLRQAOTAL 850
Db 781 TERPTGLDQAVGLPEYWEVSRKYAPFEAGLASPTGRVYHPIPGQJLSNLRQAOTAL 840
Qy 851 GLADREFELIENYAAVNMELRPTKVTSSKVVGDALHLVAGVDADFAADPOKYDIP 910
Db 841 GLGDRFEDIYAAADRMLGRVLKVTSSKVVGDALHLVAGVSPADFEQDPRDIP 900
Qy 911 DSVIAFLRGELNPPGWPPELRTALRSEGRSEKAPLTVPEEPEQAHLDADDKERNLSL 970
Db 901 DSVVGLRGELNPPGWPPELRSKALRGAEAR-PLAELSEDDRDGL----GKDRRAVL 955
Qy 971 NELLFPKTEPLEHRRFPNTSALDDREPFYGLVEGRETILRL-PDVRTPLVRLDAIS 1029
Db 956 NELLFPKTEPLEHRRFPNTSALDDREPFYGLVEGRETILRL-PDVRTPLVRLDAIS 1013
Qy 1030 EPDDKGMNRYANVNGQIRPMRVDRSVEVSTATAEKADSSNKGHVAAPFAGVVTVAE 1089
Db 1014 DADERGMRYMSSNLQRLPIQVDRSAAATDVPVTEKADRANPGHVAAPFAGVVTVAE 1073
Qy 1090 GDEVKAGDAVAIIEMKMEAITASVDGKIDRVVVVPAATKVEGGDLIV 1137
Db 1074 GDEVEAGATVATIEAMKMEASITAPKSGTVTLRAINRIQQVEGGDLIV 1121

RESULT 6
Q9F843
ID Q9F843 PRELIMINARY; PRT; 1127 AA.
AC Q9F843;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1).
GN PYC.
OS Mycobacterium smegmatis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
RN NCBI_TaxID=1772;
RP [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=20374587; PubMed=10913817;
RA Mukhopadhyay B., Purwantinil B.;
RT "Pyruvate carboxylase from Mycobacterium smegmatis: stabilization,
RT rapid purification, molecular and biochemical characterization and
RT regulation of the cellular level"; (2000).
RL Biochim. Biophys. Acta 1475:191-206 (2000).
CC -I- COPACOT: BIOTIN (BY SIMILARITY).
DR EMBL; AF262949; AAG30411.1; -.
DR HSP; P24182; 1NC.
DR GO; GO:0005737; C:cyclopasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003374; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR005479; Cbase_L_D2.
DR InterPro; IPR005481; Cbase_L_N.
DR InterPro; IPR000891; HMGL-Like.
DR InterPro; IPR003379; PYC_OADA.
DR InterPro; IPR005330; Pyruv_carbox.
DR InterPro; IPR000834; S/T_dehydrtase_BS.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF00289; Cbase_L_D2; 1.
DR Pfam; PF02786; Cbase_L_D2; 1.
DR Pfam; PF00682; HMGL-Like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRPFAM; TIGR01335; Pyruv_carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CBASE_2; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Biotin; Ligase.
SQ SEQUENCE 1127 AA; 120765 MW; 18C132C48425C67B CRC64;

Query Match 65.4%; Score 3783.5; DB 2; Length 1127;
Best Local Similarity 66.2%; Pred. No. 7.7e-193;

Matches 747; Conservative 141; Mismatches 234; Indels 7; Gaps 6;

Qy 14 KILVANRGEIAVRAAPRAALETGAATVATYPREDGRSFHRSFASAVRIGTGSPVKAYLD 73

Db 4 KLVANRGEIAVRAAPRAALETGAATVATYPREDGRSFHRSFASAVRIGTGSPVKAYLD 63

Qy 74 IDEITGAKKVADAIYGYGFLSENAQLARECAENGITGPTPEVLDLTDGKSRVTA 133

Db 64 VDEIRVAKHSGADAVYGYGFLSENPDLAKCAEAGITFVGPSAEVLQITGNKARAA 123

Qy 134 AKKAGLPVLAESTPSKNIDEIYKSAEGQTYPIFVXAVAGGGGRMRFVASPDDELKLA 193

Db 124 ARAAGLPVLSSESSESSVDELMAAADMEFFLVXAVSGGGGRMRRVTRDRESLAETEA 183

Qy 194 ASREAAAFGDAVVERAVINPQIEVQILGDHTGEVHLYERDCSLQREHKKVETAP 253

Db 184 ASREAAAFGDAVVERAVINPQIEVQILGDHTGEVHLYERDCSLQREHKKVETAP 243

Qy 254 AOHLDPDLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFTEMNPRIQVEHTVTE 313

Db 244 APNLSDELQIQICADAVAFARQIGYSCAGTVEFLDGRGHVFIENPRIQVEHTVTE 303

Qy 314 TEVDLVKQMLAAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGFRPDGTITAY 373

Db 304 TDVDLVSSQLRIAAGCTLADJLAGENTLKGALQCRITTEDPANGFRPDGTGQISAY 363

Qy 374 PGAGVRLDGAALQGGCTAHFDSMLVKMTCRGSDFETAVARACRALAEFTVSGVAT 433

```

Db      364 PGAGIRLGGTNGLGAISAHFSDMLVKLTCHGRDPSAAAGSARALAEFRIRGVSTNIP 423
QY      434 FURLAREEDFTSKRIATGFIADHPLLLOAPPADDEQGRILDYLDVTVNKPVGVRPXDV 493
Db      424 FLQAVIDDDFRRAGRVTTSFIDRRPHLLTSRSPADRGTRILNYLADITVKNPGRSPSTV 483
QY      494 AAPIDKLPNIKDL--PLPRGSRDLKQLQPAAFARDLREODALAVTDITTFDHAHQSLILAT 551
Db      484 -YPQDKLPL-LDQAPPPAGSKORVELGPEGFAGWLRKESKAVGVTDTTFDHAHQSLILAT 541
QY      552 RYRSFALKPAABAVAKLTPELLSVEWAGGATVDVAMRFLFEDPDPWRDLBELREAMPVNIQ 611
Db      542 RVRTTGLLVAPVAVRSMPLQSLIECWGATYDVALRFLKDPWRLAALRESVENICLQ 601
QY      612 MLLRGNTVGYTPYPSVCRAFVKEAASGVDFRIFFDALNDVDSQWPAIDAVLTNTAV 671
Db      602 MLLRGNTVGYTPYPLVTSAFVEEAAAGIIFRIFDALNNVESKRPALDAVRGTGTI 661
QY      672 AEVAMAYSGDLSPPNEKLYLTDYLLKVAEEIVKSGAHILAIKDMAGLRLPAPAVTKLVITAL 731
Db      662 AEVAMCYTGLSDPAENLYLTDYLLKLABQIVEAGAHVLAIKDMAGLLEAPAAHTLVSA 721
QY      732 RREFDLPVHVHTDHTAGGOLATVFAAQAQADAVDGSAPLSGTTTSQPSLSAIVAAFAHT 791
Db      722 RSRFDLPVHVHTDHTPGGQATVLAWSAGADAVDGSAPMAGTTSQPSLSAIVAAFAHT 781
QY      792 RRDITGLSLEAVSLPEYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQSLSNLRAQATALG 851
Db      782 QYDTGLDLRAVCDLEPYWEAVRKVYAPFESGLPGPTGRVYRHEIPGQSLSNLRAQATALG 841
QY      852 LADRFLIEDNYAANVEMLRPTKVTTPSSKVTGDLALHVLGAGVDPADFAADPKQYDIPD 911
Db      842 LGRDFEIEIANYAAARDVLRVLPSSKVTGDLALHVLGAGVDPADFAADPKQYDIPD 901
QY      912 SVIAFLRGLGNPPGWPPELRTRALEGRSEKAPLTEPBEQAHLDADDSKERRNSIN 971
Db      902 SVIGFLRGLGDPGWPPELRTRKALQGRPAR-PVEKLTADDEALL-AQPGPKQAALN 959
QY      972 RLIFPKPTPEFLHRRFRFGNTSLDRREFYGLVEGREFLRLPDPVTRTLLVRLDAISEP 1031
Db      960 RLIFPGPTAEFAHRETYGDTSSLSANQFYGLRYGEEHVRVQ-ERGVELLIGLEAISEA 1018
QY      1032 DDGMRNVANVANGQIRPMRVRDSVESVTATAEKADSNKGHVAAPFAGVTVTVVAEGD 1091
Db      1019 DERGMRTVMCIINGQLRPVLVRDRSIASEVPAEAKADRNADHIAAPFAGVTVTVVAEGD 1078
QY      1092 EVKAGDAVAIEAKQMEATTASVDGKIDRVVVPATKVEGGDLIVVVS 1140
Db      1079 SVDAGQTIATIEAMKEAATAPKAGTVARVAVAAATAQVEGDLIVVVS 1127

```

RESULT 7

```

P95127 PRELIMINARY; PRT; 1127 AA.
AC P95127
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Pyruvate carboxylase).
GN PCA OR RV2967C OR WTCV349.20 OR MT3045.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv.
RX MEDLINE=99295987; PubMed=9614230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Teklaia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

```

```

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544 (1998).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RC Fleischnmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RC Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RC Kolonay J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,
RC Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RC Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RP laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z83018; CAB05410.1; -;
DR EMBL; AE007125; AAK47371.1; -;
DR PIR; D70671; D70671.
DR HSSP; P24182; 1BNC.
DR TIGR; MT3045; -;
DR TubercuList; RV2967C; -;
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006520; F:amino acid metabolism; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin-carb C.
DR InterPro; IPR000899; Biotin_lipoyl.
DR InterPro; IPR005479; Cpsase_L_D2.
DR InterPro; IPR005481; Cpsase_L_N.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; PVC_OADA.
DR InterPro; IPR005930; Pyruv carbox.
DR InterPro; IPR000634; S/T dehydratase_BS.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF00289; Cpsase_L_chain; 1.
DR Pfam; PF02786; Cpsase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PVC_OADA; 1.
DR TIGRFAMs; TIGR01235; pyruv carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1127 AA; 120422 MW; 84B0A4CC1A23CD90 CRC64;

```

```

Query Match 64.1%; Score 3712.5; DB 16; Length 1127;
Best Local Similarity 64.4%; Pred. No. 4,6e-189;
Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;
QY 12 FKKLVANRGEIAYRAFAALETGAATVAIYPRDRGSHFRSFASEAVRIGTEGSPVKAY 71
Db 2 FSKVLVANRGEIATRAFAAAYELGVGTVAIYPRDRNSQHLKADSEYQIGDIGPVMAY 61
QY 72 LDIDEIIGAAKKVADATPGYGLSNAQLARECAENGTIFIGTPEVLDLTGDKSRV 131
Db 62 LSVDEIIVATARRAGADATPGYGLSENPDLLAACAAAGISFVGPSAEVLELAGNKSRAI 121
QY 132 TAAKAGLPLVLAESTPSSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRFVAPSDRLKLA 191
Db 122 AAAREAGLPVLMSGAPSASVDLLSVAAGMFPFLFVKAVAGGGGRMRFVAGDIAALPEAI 181
QY 192 TEASREAAAGDGAQVYVERAVINPQHIEVQILGDHGTGVVHYERDCSLORRQKVVEI 251
Db 182 EAASREASAFSGDPTVYLEQAVINPRHIEVQILADNLGDVHILYERDCSVORRQKVIEL 241

```

QY	252	APAQHLDELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFEMNPRIOVEHTVTE	311
Db	242	APAPHLDAELRYKCVDAVAFARHIGYSCAGTVEFLDERGEYFEMNPRVQVEHTVTE	301
QY	312	EVTEDVLKQAMRLAAGATLKGTLQDKIKTHGAALQCRITTEDPNNGPRTPTGITAY	371
Db	302	EITVDVLVASQLRAAGETLEQLGRQBDIAPHGAALQCRITTEDPANGFRPDTGRISAL	361
QY	372	RSPGGAGVRLDGAALGGEITAHFDSMLVKMTGCSDFETAVARAQALAEFTVSGVATN	431
Db	362	RTAGGAGVRLDGSNLAIEISPYFDSMLVKLTGCRDLPTAVSRARRAIAEFIRGVSTN	421
QY	432	IGFRLALLREEDFTSKRIATGFIADHPHLLQAPPADEQGRILDYLDVTVNKPVGVRPK	491
Db	422	IPFLOAVLDDPDRAGRVTTSFIDERPOLLTARASADRGTKLINFLADVTVNPNYGRPS	481
QY	492	DVAAPIDKLPNIKDLPL----PGSRDLKQLGPAFAARDLREQDALAVDTTFRDAHQ	547
Db	482	TI-YPDOKLP---DLDLRAAPPAGSKQRLVKLPGGEGFARWLKRESAAVGTDTTFRDAHQ	537
QY	548	LLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMREFLPEDPMDRLDELREAMPN	607
Db	538	LLATRVRTSGLSRVAPYLARTWPLLSEVCWGATYDVVALREFLKEDEPWERLATLRAAMPN	597
QY	608	VNIQMLLRGNTVGTTPYDPSVCRAFVKEAASSGVDI FRIIPDALNDVSMRPAIDAVLET	667
Db	598	ICLQMLLRGNTVGTTPYPEIVTSAPVQATATGIDI FRIIPDALNNISMRRPAIDAVRET	657
QY	668	NTAAVEAVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILA KDMAGLLRPAAVTKL	727
Db	658	GSAIAEVANCYTGDLTDPGEQLYTLDDYLLKLAEQIVDAGAHVLA KDMAGLLRPPAAQRL	717
QY	728	VTALRREFDLPVHVHTHTDTAGQOLATYFAAAGADAVDGSAPLSGTTQSLSAIVAA	787
Db	718	VSALRSRFDLPVHLHTHTDTPGGQLASYAAWHAGADAVDGAAPLAGTTQSALSSIVAA	777
QY	252	APAQHLDELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFEMNPRIOVEHTVTE	311
Db	242	APAPHLDAELRYKCVDAVAFARHIGYSCAGTVEFLDERGEYFEMNPRVQVEHTVTE	301
QY	312	EVTEDVLKQAMRLAAGATLKGTLQDKIKTHGAALQCRITTEDPNNGPRTPTGITAY	371
Db	302	EITVDVLVASQLRAAGETLEQLGRQBDIAPHGAALQCRITTEDPANGFRPDTGRISAL	361
QY	372	RSPGGAGVRLDGAALGGEITAHFDSMLVKMTGCSDFETAVARAQALAEFTVSGVATN	431
Db	362	RTAGGAGVRLDGSNLAIEISPYFDSMLVKLTGCRDLPTAVSRARRAIAEFIRGVSTN	421
QY	432	IGFRLALLREEDFTSKRIATGFIADHPHLLQAPPADEQGRILDYLDVTVNKPVGVRPK	491
Db	422	IPFLOAVLDDPDRAGRVTTSFIDERPOLLTARASADRGTKLINFLADVTVNPNYGRPS	481
QY	492	DVAAPIDKLPNIKDLPL----PGSRDLKQLGPAFAARDLREQDALAVDTTFRDAHQ	547
Db	482	TI-YPDOKLP---DLDLRAAPPAGSKQRLVKLPGGEGFARWLKRESAAVGTDTTFRDAHQ	537
QY	548	LLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMREFLPEDPMDRLDELREAMPN	607
Db	538	LLATRVRTSGLSRVAPYLARTWPLLSEVCWGATYDVVALREFLKEDEPWERLATLRAAMPN	597
QY	608	VNIQMLLRGNTVGTTPYDPSVCRAFVKEAASSGVDI FRIIPDALNDVSMRPAIDAVLET	667
Db	598	ICLQMLLRGNTVGTTPYPEIVTSAPVQATATGIDI FRIIPDALNNISMRRPAIDAVRET	657
QY	668	NTAAVEAVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILA KDMAGLLRPAAVTKL	727
Db	658	GSAIAEVANCYTGDLTDPGEQLYTLDDYLLKLAEQIVDAGAHVLA KDMAGLLRPPAAQRL	717
QY	728	VTALRREFDLPVHVHTHTDTAGQOLATYFAAAGADAVDGSAPLSGTTQSLSAIVAA	787
Db	718	VSALRSRFDLPVHLHTHTDTPGGQLASYAAWHAGADAVDGAAPLAGTTQSALSSIVAA	777
QY	252	APAQHLDELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFEMNPRIOVEHTVTE	311
Db	242	APAPHLDAELRYKCVDAVAFARHIGYSCAGTVEFLDERGEYFEMNPRVQVEHTVTE	301
QY	312	EVTEDVLKQAMRLAAGATLKGTLQDKIKTHGAALQCRITTEDPNNGPRTPTGITAY	371
Db	302	EITVDVLVASQLRAAGETLEQLGRQBDIAPHGAALQCRITTEDPANGFRPDTGRISAL	361
QY	372	RSPGGAGVRLDGAALGGEITAHFDSMLVKMTGCSDFETAVARAQALAEFTVSGVATN	431
Db	362	RTAGGAGVRLDGSNLAIEISPYFDSMLVKLTGCRDLPTAVSRARRAIAEFIRGVSTN	421
QY	432	IGFRLALLREEDFTSKRIATGFIADHPHLLQAPPADEQGRILDYLDVTVNKPVGVRPK	491
Db	422	IPFLOAVLDDPDRAGRVTTSFIDERPOLLTARASADRGTKLINFLADVTVNPNYGRPS	481
QY	492	DVAAPIDKLPNIKDLPL----PGSRDLKQLGPAFAARDLREQDALAVDTTFRDAHQ	547
Db	482	TI-YPDOKLP---DLDLRAAPPAGSKQRLVKLPGGEGFARWLKRESAAVGTDTTFRDAHQ	537
QY	548	LLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMREFLPEDPMDRLDELREAMPN	607
Db	538	LLATRVRTSGLSRVAPYLARTWPLLSEVCWGATYDVVALREFLKEDEPWERLATLRAAMPN	597
QY	608	VNIQMLLRGNTVGTTPYDPSVCRAFVKEAASSGVDI FRIIPDALNDVSMRPAIDAVLET	667
Db	598	ICLQMLLRGNTVGTTPYPEIVTSAPVQATATGIDI FRIIPDALNNISMRRPAIDAVRET	657
QY	668	NTAAVEAVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILA KDMAGLLRPAAVTKL	727
Db	658	GSAIAEVANCYTGDLTDPGEQLYTLDDYLLKLAEQIVDAGAHVLA KDMAGLLRPPAAQRL	717
QY	728	VTALRREFDLPVHVHTHTDTAGQOLATYFAAAGADAVDGSAPLSGTTQSLSAIVAA	787
Db	718	VSALRSRFDLPVHLHTHTDTPGGQLASYAAWHAGADAVDGAAPLAGTTQSALSSIVAA	777
QY	252	APAQHLDELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFEMNPRIOVEHTVTE	311
Db	242	APAPHLDAELRYKCVDAVAFARHIGYSCAGTVEFLDERGEYFEMNPRVQVEHTVTE	301
QY	312	EVTEDVLKQAMRLAAGATLKGTLQDKIKTHGAALQCRITTEDPNNGPRTPTGITAY	371
Db	302	EITVDVLVASQLRAAGETLEQLGRQBDIAPHGAALQCRITTEDPANGFRPDTGRISAL	361
QY	372	RSPGGAGVRLDGAALGGEITAHFDSMLVKMTGCSDFETAVARAQALAEFTVSGVATN	431
Db	362	RTAGGAGVRLDGSNLAIEISPYFDSMLVKLTGCRDLPTAVSRARRAIAEFIRGVSTN	421
QY	432	IGFRLALLREEDFTSKRIATGFIADHPHLLQAPPADEQGRILDYLDVTVNKPVGVRPK	491
Db	422	IPFLOAVLDDPDRAGRVTTSFIDERPOLLTARASADRGTKLINFLADVTVNPNYGRPS	481
QY	492	DVAAPIDKLPNIKDLPL----PGSRDLKQLGPAFAARDLREQDALAVDTTFRDAHQ	547
Db	482	TI-YPDOKLP---DLDLRAAPPAGSKQRLVKLPGGEGFARWLKRESAAVGTDTTFRDAHQ	537
QY	548	LLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMREFLPEDPMDRLDELREAMPN	607
Db	538	LLATRVRTSGLSRVAPYLARTWPLLSEVCWGATYDVVALREFLKEDEPWERLATLRAAMPN	597
QY	608	VNIQMLLRGNTVGTTPYDPSVCRAFVKEAASSGVDI FRIIPDALNDVSMRPAIDAVLET	667
Db	598	ICLQMLLRGNTVGTTPYPEIVTSAPVQATATGIDI FRIIPDALNNISMRRPAIDAVRET	657
QY	668	NTAAVEAVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILA KDMAGLLRPAAVTKL	727
Db	658	GSAIAEVANCYTGDLTDPGEQLYTLDDYLLKLAEQIVDAGAHVLA KDMAGLLRPPAAQRL	717
QY	728	VTALRREFDLPVHVHTHTDTAGQOLATYFAAAGADAVDGSAPLSGTTQSLSAIVAA	787
Db	718	VSALRSRFDLPVHLHTHTDTPGGQLASYAAWHAGADAVDGAAPLAGTTQSALSSIVAA	777

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1765;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Gartner T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
 Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon P.,
 Harris B., Atkin R., Doggett J., Mayes R., Keating T., Wheeler S.R.,
 Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 DR EMBL; BX248344; CAD96678.1; --
 KW Ligase; Complete proteome.
 SQ SEQUENCE 1127 AA; 120423 MW; 84B0A4CC1A23CD90 CRC64;

Query Match 64.1%; Score 3712.5; DB 16; Length 1127;
 Best Local Similarity 64.4%; Pred. No. 4.6e-189;
 Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;
 QY 12 FKKILVANRGEIATVAFRAALETGAATVAIYPREDRGSFHRSEAFVRIETGSPVKAY 71
 Db 2 FSKVLVANRGEIATVAFRAAYELGVGTVAIVPYEDRNSQHLKADESQIGDIGHPVHAY 61
 *QY 72 LDIDIIIGAAKKVADAIYPGYFLSENAQLARCAENGITFIPTPEVLDITGDKSEAV 131
 Db 62 LSVDEIVATARRAGADAIYPGYFLSENPDLAAACAAAGISFVGPSAEVLEAGNKSRAI 121
 QY 132 TAACKAGIPLVAESPPSKNIDIVKASGQYPIFVKAVAGGGGRGMFVSPDELRKLA 191
 Db 122 AARERAGIPLVMSAPSASVDELLSVAAGMFPFLFVKAVAGGGGRGMFVSPDELRKLA 181
 QY 192 TFASEAAAFGAGVAVVERAVINPQHIEVOILGDHTGEVHLHYERDCSLQRHKKVVEI 251
 Db 182 EAASREASAFGDPVYVLEQAVINPRHIEVQILADNLGDVHLHYERDCSVQRHKKVIEL 241
 QY 252 APAQHLDELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFEMNPRIOVEHTVTE 311
 Db 242 APAPHLDAELRYKCVDAVAFARHIGYSCAGTVEFLDERGEYFEMNPRVQVEHTVTE 301
 QY 312 EYTEVDVLKQAMRLAAGATLKGTLQDKIKTHGAALQCRITTEDPNNGPRTPTGITAY 371
 Db 302 EITVDVLVASQLRAAGETLEQLGRQBDIAPHGAALQCRITTEDPANGFRPDTGRISAL 361
 QY 372 RSPGGAGVRLDGAALGGEITAHFDSMLVKMTGCSDFETAVARAQALAEFTVSGVATN 431
 Db 362 RTAGGAGVRLDGSNLAIEISPYFDSMLVKLTGCRDLPTAVSRARRAIAEFIRGVSTN 421
 QY 432 IGFRLALLREEDFTSKRIATGFIADHPHLLQAPPADEQGRILDYLDVTVNKPVGVRPK 491
 Db 422 IPFLOAVLDDPDRAGRVTTSFIDERPOLLTARASADRGTKLINFLADVTVNPNYGRPS 481
 QY 492 DVAAPIDKLPNIKDLPL----PGSRDLKQLGPAFAARDLREQDALAVDTTFRDAHQ 547
 Db 482 TI-YPDOKLP---DLDLRAAPPAGSKQRLVKLPGGEGFARWLKRESAAVGTDTTFRDAHQ 537
 QY 548 LLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMREFLPEDPMDRLDELREAMPN 607
 Db 538 LLATRVRTSGLSRVAPYLARTWPLLSEVCWGATYDVVALREFLKEDEPWERLATLRAAMPN 597
 QY 608 VNIQMLLRGNTVGTTPYDPSVCRAFVKEAASSGVDI FRIIPDALNDVSMRPAIDAVLET 667
 Db 598 ICLQMLLRGNTVGTTPYPEIVTSAPVQATATGIDI FRIIPDALNNISMRRPAIDAVRET 657
 QY 668 NTAAVEAVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILA KDMAGLLRPAAVTKL 727
 Db 658 GSAIAEVANCYTGDLTDPGEQLYTLDDYLLKLAEQIVDAGAHVLA KDMAGLLRPPAAQRL 717
 QY 728 VTALRREFDLPVHVHTHTDTAGQOLATYFAAAGADAVDGSAPLSGTTQSLSAIVAA 787
 Db 718 VSALRSRFDLPVHLHTHTDTPGGQLASYAAWHAGADAVDGAAPLAGTTQSALSSIVAA 777

RESULT 8
 Q7TXJ1
 ID Q7TXJ1 PRELIMINARY; PRT; 1127 AA.
 AC Q7TXJ1;
 DT 01-OCT-2003 (TreeBLrel. 25, Created)
 DT 01-OCT-2003 (TreeBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TreeBLrel. 25, Last annotation update)
 DE Probable pyruvate carboxylase PCA (Pyruvic carboxylase)
 DE (EC 6.4.1.1).
 GN PCA OR MB2991C.
 OS Mycobacterium bovis.

```
QY 788 FAHTRDTGLSLAEVSDLEPYWEAVRGYLYLPFESGTPGTGRVYRHEIPGQGLSNRAQA 847
Db 778 AAHTEDYDGLSLAYCALEPYWEALRYKYPFESGLPGFTGRVYRHEIPGQGLSNRQQA 837
QY 848 TALGLADREFELIEDYAAVNMELGRPTKVTPTSSKVVGDALHLVAGVDPADFAADPKQY 907
Db 838 IALGLDGRFERIEEYAGADRVGLRVKVTPTSKVVGDALALVAGVSADEFASDPARP 897
QY 908 DIPDSVIAFLRGEINPGGWPPEPLRTALGRSEKAPLTVPEEEOAHLDADDKERR 967
Db 898 GIPESVGLFRLGELGDPGPGWPEPLRTAALAGRGAAR-PTAQLAADBEIALSSVGAK-RQ 955
QY 968 NLGNLLPKPTEEFLEHRRRNGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1027
Db 956 ATLAKLLFPSPKFEFNEHREAGDTSQLSANQFFYGLRQGEHRVKL-ERGVELLIGLEA 1014
QY 1028 ISEPDGKMRNVANVNGOIRPMVRDRSVSVTATAKADSSNKGHVAAPFAGVVTVV 1087
Db 1015 ISEPDGMRVTWCINGLRVLRDRSIAASAVPAEAKADRGNFGHTAAPFAGVVTVG 1074
QY 1088 ASGDEKAGDAVATTEAMKWEATITASVDGKIDRVVPAATKVEGGDLVVVS 1140
Db 1075 CVGERVAGQTIATTEAMKWEAPITAPVAGTVVERVAVSDTAQVEGGDLLVVVS 1127

RESULT 9
Q50450
ID Q50450 PRELIMINARY; PRT; 1124 AA.
AC Q50450;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pyc.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID:1773;
RN [1]
RA Smith D.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; U00024; AAA50948.1; -.
DR HSSP; P24182; 1BNC.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPRO01882; Biotin_BS.
DR InterPro; IPRO05482; Biotin_Carb_C.
DR InterPro; IPRO00089; Biotin_lipoyl.
DR InterPro; IPRO05479; Cphase_L_D2.
DR InterPro; IPRO05481; Cphase_L_N.
DR InterPro; IPRO00891; HMGL-Like.
DR InterPro; IPRO03379; Pyc OADA.
DR InterPro; IPRO05930; Pyluv carbox.
DR InterPro; IPRO00634; S/T dehydrtse_BS.
DR Pfam; PF02785; Biotin_cab_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; Cphase_L_chain; 1.
DR Pfam; PF02786; Cphase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; Pyc OADA; 1.
DR TIGRfams; TIGR01235; Pyluv carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
```

```
DR PROSITE; PS00867; CPSASE 2; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Biotin.
SQ SEQUENCE 1124 AA; 120672 MW; 07E1692E12203491 CRC64;
Query Match 60.2%; Score 3483; DB 2; Length 1124;
Best Local Similarity 62.2%; Pred. No. 7.6e-177;
Matches 698; Conservative 151; Mismatches 251; Indels 22; Gaps 9;
QY 12 FKILVANGETAVAFRAALETGTAATVAYIPREORGSPHRSFASAEVRIGTEGSPVKAY 71
Db 2 FSKVLVANGETAIFRAAYELGVGTVAIVPYEDRNSOHLKADSEYIGIDGHPVHAY 61
QY 72 LDDEIIGAIAKKVADAIYPGYGFLSENAQALARECAENGITIFGPTPEVLDITGDKSRAV 131
Db 62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGSAEVLGAGNKSRAI 121
QY 132 TAAKKAGLPVLAESTPSKNIIDEIVKSABEQTPVIFKAVAGGGGRMRVAFSPDELKLA 191
Db 122 AAAREAGLPVLMSSAFSASVDELLSVAAGMPFELFKAVAGGGGRMRVAGDIALPEAI 181
QY 192 TEASREAAAFGDCAVYVERAVINPOHIEVQILGHTGSEVHLYERDCSLORRHQKVVEI 251
Db 182 EASRREASAFGDPVYLEQAVINPHEIEVQILADNLGDIHLYERDCSVQRHOKVIEL 241
QY 252 APAQHLDPELRDRI CADAVKFCRSIOYQAGTVFEFLVDEKGNHVFIEHMPRQVETVTE 311
Db 242 APAPHLDAELRYKMCVDVAFARHIGYCAGTVFEFLDERGEYVFIEHMPRQVETVTE 301
QY 312 EVTEVDLVKAQMRLLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGITAY 371
Db 302 EITDVLVASQRIAGATEILEQLGQEDIAHPGALQCRITTEDPANGFRP-TRAGSAR 360
QY 372 RSPGGAGVRLDGAALQGGIEITHAFDMLVKM-----TCRGSFETAVARAQALAEFTVS 426
Db 361 CDPPAVPVSAWTAAP-----TWERNQVFLRLHAGAQADLSGRDLPTAVSARRAIAEFRIR 415
QY 427 GVAITNIGFLRALLREEDFTSKRTATGFIADHPHLLQAPPADDEQGRILDVLADVTNKKPH 496
Db 416 GVSTNIPFLQAVLDLDDDFRAGRVTTSFIDERPOLLTARASADRGTKILNPLADVTNNFY 475
QY 487 GVRPKQVAAPIDKLPNIKOLF-----PRGSRDLKQLGPAAPARDLREQDALAVTTTFR 542
Db 476 GSRPSTI-YPDCKLP---DLDLRAAPPAGSKQLVKLGPEGFARWLRESAAVGTDTTFR 531
QY 543 DAHQSLLATRVRSFALKPAEAEVAKLTPELLSVEAMGGATYDVAMFLPEDPMDRLDELRL 602
Db 532 DAHQSLLATRVRTSGLSRVAPYLARTWMPQLLSVECGGATYDVVALRFLKEDPHERLATLR 591
QY 603 EAMPNVNIQMLLRGRNTVGYTPYPSVCRAFVKAEASGVDIIFRIFDALNDVDSQMRPAID 662
Db 592 AAMENICLQMLLRGRNTVGYTPYPEIVTSAFVQEAATATGIDIIFRIFDALNNIESMRPAID 651
QY 663 AVLETNTAVAEVAMVAGSDLSDENEKLYTLDYVYKMAEETVKSGAHILATKDWAGLLRPA 722
Db 652 AVRETGSAIAEVAMCVTGDLTDPGEQLYLDYVYKLAEQIVDAGAHVLAIKDWAGLLRPP 711
QY 723 AVTKLVTALRREFDLPVHVHTHTDTAGQLATYFAAAQAGADAVDGSAPLSGTTSPSLS 782
Db 712 AAQELVSALRSRFDLPVHLHTHTDTPGQLASVYAAWHAGADAVDGAAPLAGTTSQPALS 771
QY 783 AIVAAFAHTRDRDTGLSLEAVSDLEPYWEAVRGYLYLPFESGTPGTGRVYRHEIPGQGLSN 842
Db 772 SIIVAAAHAETEYDTGLSLSAVCALEPYWEALRYKYPFESGLPGFTGRVYRHEIPGQGLSN 831
QY 843 LRAQATALGLADREFELIEDYAAVNMELGRPTKVTPTSSKVVGDALHLVAGVDPADFAA 902
Db 832 LRQQAIALGLCDRFEIEEYAGADRVGLRVKVTPTSKVVGDALALVAGVSADEFAS 891
QY 903 DPQKYDIPDSVIAFLRGEINPGGWPPEPLRTALGRSEKAPLTVPEEEOAHLDADD 962
Db 892 DPAREFGIPESVLGFLRGEINPGGWPPEPLRTAALAGRGAAR-PTAQLAADBEIALSSVG 950
```

QY	963	SKERNLSNRLLPFKPTSEFLEHRRRFRONTSSALDRREFFVGLVEGRETLIRLPDVTPLL	1022
Db	951	AK-ROATLNRLLPSPFTKFEHREAYGDTSQLSANOFFYGLRQGEHRVK--ERGVELL	1008
QY	1023	VRLDAISBDDKGNNVNANVGRIPMVRVDRAVESVTATAEKADSSNKGHVAAPFAGV	1082
Db	1009	IGBAISEPDDEGRMTVMCIINGQLRPVLVLDRIAASAVPAEKAADRGNHGHAAPFAGV	1068
QY	1083	VTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVV	1124
Db	1069	VTVGVCGVERVAGGTIATIEAMKMEAFITPAGVTVERNAV	1110
RESULT 10			
O83HF3	PRELIMINARY; PRT: 1131 AA.		
ID	Q83HF3		
AC	Q83HF3;		
DT	01-JUN-2003 (TrEMBLrel. 24, Created)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)		
DE	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Pyruvate carboxylase (EC 6.4.1.1).		
GN	PyC OR TW647.		
OS	Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
CC	Micrococineae; Cellulomonadaceae; Tropheryma.		
OX	NCBI_TaxID=218496;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22495039; PubMed=12606174;		
RA	Bentley S.D., Maitland M., Murphy L.D., Pallen M.J., Yeats C.A.,		
RA	Dover L.G., Norbertczak H.T., Besta G.S., Quail M.A., Harris D.E.,		
RA	von Herbay A., Goble A., Rutter S., Squares R., Squares S.,		
RA	Barrell B.G., Parkhill J., Rellman D.A.;		
RT	"Sequencing and analysis of the genome of the Whipple's disease		
RT	bacterium Tropheryma whipplei.";		
RL	Lancet 361:637-644(2003).		
ENBL	BX251412; CAD67310.1;		
DR	GO; GO:0005737; C:citoplasm; IEA.		
DR	GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:000374; F:biotin binding; IEA.		
DR	GO; GO:0016874; F:lignase activity; IEA.		
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR001969; Aspprotease_A8.		
DR	InterPro; IPR001882; Biotin_BS.		
DR	InterPro; IPR005482; Biotin_Carb.C.		
DR	InterPro; IPR000085; Biotin_Lipoyl.		
DR	InterPro; IPR005479; Cpase_L_D2.		
DR	InterPro; IPR005481; HMGL-Like.		
DR	InterPro; IPR003379; PyC_OADA.		
DR	pfam; PF02785; Biotin_Carb.C. 1.		
DR	pam; PF00364; biotin_lipoyl. 1.		
DR	pam; PF00289; CPSase_L_chain. 1.		
DR	pam; PF02786; CPSase_L_22. 1.		
DR	pam; PF00682; HMGL-like. 1.		
DR	pfam; PF02436; PyC_OADA. 1.		
DR	TIGRFAMS; TIGR01235; pyruv carbox. 1.		
DR	PROSITE; PS00141; ASP_PROTEASE. 1.		
DR	PROSITE; PS00188; BIOTIN. 1.		
DR	PROSITE; PS00867; CPSASE.2. 1.		
KW	Ligase; Complete proteome.		
SQ	SEQUENCE 1131 AA; 124343 MW; 6940857BF7A3A981 CRC64;		
Query Match 56.1%; Score 3248; DB 16; Length 1131;			
Best Local Similarity 58.4%; Pred.No.2.se-164;			
Matches 669; Conservative 152; Mismatches 290; Indels 34; Gaps 10;			
OY	12	FKKILVANRGEIAVRAPFAALETKAATVAIYPREDRGSHRFSFAGEAVRICGTGSPVKAY	71

Db 1065 IASPPAGQVTKVDVGDEVVSGQAVALEAMKMTTVNAPVSGQVIRISIPFGQVDIGD 1124

Qy 1135 LIVVV 1139

Db 1125 LIMEI 1129

RESULT 11

Q83FSS PRELIMINARY; PRT; 1131 AA.

AC Q83FSS

DT 01-JUN-2003 (T-EMBLrel. 24, Created)

DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Pyruvate carboxylase [EC 6.4.1.1].

GN PCA OR TW630.

OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Micrococccineae; Cellulomonadaceae; Tropheryma.

OX NCBI_TaxID=203267;

RN [1]

RP SEQUENCE FROM N.A.

RA Raoult D., Audic S., Robert C., Ogata H., Suhr K., Drancourt M.,

RA Claverie J.-M.;

RT "Tropheryma whipplei illustrates the diversity of gene loss patterns

in small genome bacterial pathogens";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE01852; AA04472.1; -

DR GO; GO:0005737; C:cytoplasm; IEA.

DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0009374; F:biotin binding; IEA.

DR GO; GO:0016874; F:ligase activity; IEA.

DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.

DR GO; GO:0006094; P:gluconeogenesis; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001969; Asprotease_AS.

DR InterPro; IPR001882; Biotin_BS.

DR InterPro; IPR005482; Biotin_carb_C.

DR InterPro; IPR000089; Biotin_lipoYL.

DR InterPro; IPR005479; CPase_L_D2.

DR InterPro; IPR005481; CPase_L_N.

DR InterPro; IPR000891; HMGL-like.

DR InterPro; IPR003379; PYC_OADA.

DR InterPro; IPR005930; Pyruv_carbox.

DR Pfam; PF02785; Biotin_carb_C; 1.

DR Pfam; PF00364; biotin_lipoYL; 1.

DR Pfam; PF00289; CPase_L_chain; 1.

DR Pfam; PF02786; CPase_L_D2; 1.

DR Pfam; PF00682; HMGL-like; 1.

DR Pfam; PF02436; PYC_OADA; 1.

DR TIGRfams; TIGR01235; pyruv_carbox; 1.

DR PROSITE; PS00141; ASP_PROTEASE; 1.

DR PROSITE; PS00188; BIOTIN; 1.

DR PROSITE; PS00867; CPASE_2; 1.

KW Ligase; Complete proteome.

SQ SEQUENCE 1131 AA; 124300 MW; 6C21F973DA524CB1 CRC64;

Query Match 55.0%; Score 3241; DB 16; Length 1131;

Best Local Similarity 58.3%; Pred. No. 5.9e-164;

Matches 668; Conservative 152; Mismatches 291; Indels 34; Gaps 10;

Qy 12 FKILVANRGEIARAFRAALETGAATVAIYPREDRGSFHSRSPASEAVRIGTSGSPKAY 71

Db 2 FKILIANRGEIARIRGRAAPERQIQTVAIYAHEDRNSLRKADSBAYQIGQGSPPVAA 61

Qy 72 LDIDEIIIAAKKADAIYGYGLSENAGLARECAENGITFTICPTPEVLDTGDKSRV 131

Db 62 LDVHEILIRVALMSNSDAIHFGYGLSSEYLLADEAEKNGITFTICPTPEVLDTGDKSRV 121

Qy 132 TAAKAGLPVLAESTPKNIDEIVKSGEQTYPIFVAVAGGGGRGMRFVASPDELRKLA 191

RESULT 12

Q8UHF6

Db 122 HMAQAAGLPTLRSTASSNYDELLREAELEFYFVKAAASGGGGRGMRIVENTATKNSL 181

Qy 192 TEASREAEAFGDCAVYVRAVINPOHIEVQILGCDHTGEVVHLYYERDCSLQRHOKVVEI 251

Db 182 ESAIOEAAASFGDPRVLEFETALDKPRHIEVQVLADKEGNIHLFERDCSLQRHOKVIEI 241

Qy 252 APAQHLDPRLDRICADAVKFCRSIGYQAGATVEFLVDEKGNHVFIEMNRIQVEHVTVE 311

Db 242 APAFNIPELLRTTLYRDAIAFAKSVKVENAGTVEFLVDSKONHYFIEMNRIQVEHVTVE 301

Qy 312 EVTEVDLVKAQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGITAY 371

Db 302 EITDIDIVOSIILIAAGASLDDIGLVODKIERRGFALQCRITTEDPHANFRPDIGRTSY 361

Qy 372 RSPGAGVRLDGAA-QJGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 430

Db 362 QSPGAGIRLDASAVNFCVEITPYFDSMLVKMTCRGSFADATNRAERGLAEFRVGVAT 421

Qy 431 NIGFLRALLREEDFTSKRIATGFTADPHLLQAPPADDEQGRILDLVADTVNKPGRVP 490

Db 422 NISFLRTLLDSSNFLNADFTTFIEKNGYLLQGSFLDKHDLVSYLGHTVVKPYGDRP 481

Qy 491 KDVAAPIKIDPNIKDLPLP-----RGRDRLLKOLGPAFAFARDREQDALAVTDTTF 541

Db 482 -ELVDPCKSIANF---LPDSSEVKSVESGRDVLRLGPGFAKWLCCRGLAVTDTTF 536

Qy 542 RDAHQSLATLTVRSFALKPAEAAVAKLTPELLSVKMGATYDVAMFLPEDPWRDLDEL 601

Db 537 RDAHQSLATLTVRTIDLSRAEACTSAALPELFSVEVGGATYDVALFLVEDPHERLSKI 596

Qy 602 REAMPNVNIQMLRGRNTVGYTPYDSCVRAFYKAAASSGVDIPIFDALNDVSQMRPAI 661

Db 597 REKVGSIQLOMLRGRNTVGYTPYDQVTRAFVDEASDLGIDIPRIFDALNDVDQMRIAI 656

Qy 662 DAVLETTTAAVAEAVMAYSGDLSDPNEKLYLDVYLKVAEEIVKSGAHLAIKDWAGLLRP 721

Db 657 DAVQQTN-SVAEVAICYGDDLLKREVTYIDYLEAKKIVDGAHILAIKDWAGVLRP 715

Qy 722 AAVTKLVTALRRFELPVHVTHTDGTAGQLATYFAAAQAGADAVDGSAPLSGTTSQPSL 781

Db 716 RAATLLVSALKREFALPVHLTHDTDPGQLATILAAADSGYDAVDVASGPMSTTSQPSM 775

Qy 782 SAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGILYLPESGTPGTPGVYHEHPGGOLS 841

Db 776 SSLVAATDNTHEHTGLSLSRVNELEPYWEAVRRLYVPFESGLLSTGRTGVYIHEIPGGOLS 835

Qy 842 NLRAQATALGLADREFELTDNYAAVNEMLGRFTKVTPSSKVVYVGLALHLVAGVDPADFA 901

Db 836 NLKQQAIALGLSDRFEIIEEMVAVYNTLFGRI PKVTPSSKVVYVGLALVLAWSNPDLGDFE 895

Qy 902 ADPOKYDIPDSVIAFLRGLGNPFGGWPEPLRTALGRSEKAPLFEVPEERQAHLDAD 961

Db 896 MNPKKYDIPDSVIFLAGELGTPPAGWPD-FRDRVLAER-----EI-SIEQHPJSSD 945

Qy 962 DSKE-----RRNSLNRLLPKPTTEFLEHRRFNGNTSALDDREFFYGLVEGRETJRL 1014

Db 946 DSKNLATSGKIRQQTLSKLEPEPYRAFEANRAEYGLDLSLRSEFPVGLDFGIEYKIAV 1005

Qy 1015 PDVTEPLLVRDLDAISEPDDKGRNVANVNGQIIRPMVRDRSVESVTATARKAASSNKGH 1074

Db 1006 SS-SVGILVRLAEIAGVDSKGRSLVLSVNGELRPQIVRDESANVVEVSRAEKADPNPFGH 1064

Qy 1075 VAAPAGVVTVAEGDEVKAGDAVATIEAMKMEATITASVDGKIDRVVVAATKVEGGD 1134

Db 1065 IASPPAGQVTKVDVGDEVVSGQAVALEAMKMTTVNAPVSGQVIRISIPFGQVDIGD 1124

Qy 1135 LIVVV 1139

Db 1125 LIMEI 1129

ID	Q8JHF6	PRELIMINARY;	PRT; 1178 AA.	
AC	Q8JHF6;			
DT	01-OCT-2002 (Tremblrel. 22, Created)			
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	Pyruvate carboxylase (EC 6.4.1.1).			
GN	PYC.			
GN	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22145742; PubMed=12150961;			
RA	Jirapakdee S., Nezc M.G., Cassidy A.I., Khew-Goodall Y.,			
RA	Wallace J.C.;			
RT	"Molecular cloning and domain structure of chicken pyruvate			
RT	carboxylase."			
RL	Biochem. Biophys. Res. Commun. 295:387-393 (2002).			
DR	EMBL; AF509529; AAM92771.1;			
DR	GO; GO:0005737; C:cytoplasm; IEA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0036874; F:ligase activity; IEA.			
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.			
DR	GO; GO:0006094; P:metabolism; IEA.			
DR	GO; GO:0008152; P:gluconeogenesis; IEA.			
DR	InterPro; IPR000089; Biotin carb C.			
DR	InterPro; IPR000482; Biotin lipoyl.			
DR	InterPro; IPR005479; Cysase L D2.			
DR	InterPro; IPR005481; Cysase L N.			
DR	InterPro; IPR000891; HMG-Like.			
DR	InterPro; IPR003379; PYC_OADA.			
DR	InterPro; IPR005930; Pyruv. carbox.			
DR	Pfam; PF02785; Biotin carb C; 1.			
DR	Pfam; PF00364; Biotin lipoyl; 1.			
DR	Pfam; PF00289; Cysase L chain; 1.			
DR	Pfam; PF02786; Cysase L D2; 1.			
DR	Pfam; PF00682; HMG-Like; 1.			
DR	Pfam; PF02436; PYC_OADA; 1.			
DR	TIGRFAms; TIGR01235; pyruv. carbox; 1.			
KW	Ligase.			
SQ	SEQUENCE 1178 AA; 127253 MW; F0722FEAB8BF39A5 CRC64;			
	Query Match 44.1%; Score 2552; DB 13; Length 1178;			
	Best Local Similarity 48.1%; Pred No. 3, 1e-127;			
	Matches 556; Conservative 165; Mismatches 410; Indels 26; Gaps 14;			
QY	2 STHSTSLPAPKILVANRGEIAVRAALETGAATVAIYIPREDRGSFHRSPASEAVRI 61			
Db	28 SVRSACQP-IRKLVANRGEIAVRAALETGAATVAIYIPREDRGSFHRSPASEAVRI 86			
QY	62 GTEGSPVKAYLDIDETIGAAKKVADAIYPGYFLSENAQLARECAENGITFIPTPEVL 121			
Db	87 GRGLPPVQALHVPDIIRVARENAVDALHPGFLSERADFAQACVDAVGFVGFPEVV 146			
QY	122 DLTDGKSRVATAAKAGLPVL-AESTPSKNIDIEIVKSAEGQTYPIFVKAAGGGGRNRF 180			
Db	147 RWMGDKVARSIAIAGVVPVPTGAPVATLGEAQDFAARVGFPIIFKAARHGGRGMRA 206			
QY	181 VASPELRLKATEASREAAAGDGVYVERAVINPOHTEVQILGDHTGEVHLYERDCS 240			
Db	207 VRGPQLEESFGRASEALAFDGGALFVEKLMERPHIEGQILGDKHGNVHLYERDCS 266			
QY	241 LQRHKKVVEIAPAQHLDPDLADRICADAVKFCRSIGVQAGTVEFLVDEKGNHVFIEMN 300			
Db	267 IQRHKKVVEIAPAQHLDPDLADRICADAVKFCRSIGVQAGTVEFLVDEKGNHVFIEMN 326			
QY	301 PRIOVHTVTEVTEVDIVKAQMLAAGATLKEIGLTQDKIKTHGAALOCRTTDPNNG 360			
Db	327 SRLQVHTVTEITGTVGLVQALVAAGRSLSGLQDQSVNAGCAICQRTTDPNNG 386			
QY	361 FRPDTGTITAYRSPGAGVRLDGAQL-GGEITAHFDSMLVMTCRGSDFFETAVARAQRA 419			

Db	387 PQPDTGRTEVERSGEGMGRILDGASAFQCALISPHYDSLIVKVIAHGPDQPSAAAKSRA 446			
QY	420 LAEFTVSVATNIGFLRALLREEDFTSKRIATGFIADHPLHAPPADDEQCRILDYAD 479			
Db	447 LGEFRIRGKTNIPFLQNVLAHPQLGGAVDTQIDENFELFHLRPSQRAKLLHLIGH 506			
QY	480 VIVNPKPHGVRP-KDVAAPIDKLPNIKOLPLPRGS-----RDLKOLGPAAPARDLRQDA 533			
Db	507 VMVNGPSTPLPVKAKAAVVEFVPP-----PVMGSPPEGLRAVLQREGPAGFARALGRHG 562			
QY	534 LAVTDTTFRDAHQSLATFVRSFALKPAAEAVAKLTPELLSVEAWGATYVAMRFIFED 593			
Db	563 LLLXDTTFRDAHQSLATFVRSFALKPAAEAVAKLTPELLSVEAWGATYVAMRFIFED 622			
QY	594 PWDRLDELREAMPVNIQMLLRGNVTGYTPYPSVCSAFVKEAASSGVDIFRIFDALND 653			
Db	623 PWERLRELRLVFNIPFQMLLRGANVGYTNPVNIYRFEVAAANGMDIFRIFDALNY 682			
QY	654 VSQMPADDAVLETTAVAEVAMAYSGDLPNEKLYTLDVYLKMAEIVKSGAHILAIAK 713			
Db	683 LPNILLGVEAVGRAG-AVVEAALSYTGVDVPTTKYSLDYLGLAKELVAGTHILCIK 741			
QY	714 DMAGLLRPAAVTKLVTLRRBF-DLPVHVHTHTAGGQATYFAAAQAGADAVDGASAPL 772			
Db	742 DMAGLLTFAAEALLVSSLRDRPDPVPIHVHTHTAGAAATLLAAANADADVDVAVDAM 801			
QY	773 SGTTSQPSLSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGR--V 830			
Db	802 SGMTSQPSMGALVACARGTPLDTGIALRVEFYSEYEWGARGLYAAFDCTATMKSNGADV 861			
QY	831 YRHEITPGGQSLNRAQATAGLADRFEITEDNYAAVNEMLGRPTKVTPESSKVVGDALHL 890			
Db	862 YENETPGGYTNLHFQAHMGLGHKFKYAKAYAEANKLLGDLIKVTPSSKVVGDALQFM 921			
QY	891 VGAGVDPAADADPOKIDIPDSVIAFLRGELGNPPGHPPELRTALLEG--RSEKAPLT 948			
Db	922 VONGLSREAEAREADELSFPLSVFVFLQYIGTPPGPPEPPRPSKVLKDLRVEGRPGAS 981			
QY	949 EYPEBEOA---HLDDADS--KERRNSLNRLFPKTEBFEHRRRFGNTSALDDREFFYG 1003			
Db	982 LPPLDFEALSQELGARDGTPSPEDLLSALYPKVYAEFRDFTSTFGPVSCIGTRIFLEG 1041			
QY	1004 LVEGRETILRLPDVTRTFLVRLDAISEPDDKMRNVVANVQIRPMVRDRSVESVAT 1063			
Db	1042 PTIAEEFEVELERGT-LHIKALGDLNAAQREAFPELNCQLRSILVRDQALKEHV 1100			
QY	1064 AEKADSSNKGHVAAPAG-VVTVTVAGDEVKAGDAVAIIEMKQEAATITASVDGKIDRV 1122			
Db	1101 HPKADRSAGQVAGPMPGEVVEVRVKEGAEKGAFLCVLSAMKMETVTVAPRGGTVSEL 1160			
QY	1123 VVPAATKVEGGDLIVVV 1139			
Db	1161 HVRFPGMSLEGDDLIABI 1177			

RESULT 13

ID	Q7YS28	PRELIMINARY;	PRT; 1178 AA.	
AC	Q7YS28;			
DT	01-OCT-2003 (Tremblrel. 25, Created)			
DT	01-OCT-2003 (Tremblrel. 25, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	Pyruvate carboxylase.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	Tissue=Liver;			
RA	Davalos G., Sanchez A., Noguera J.L., Amills M.;			
RT	"Nucleotide sequence, polymorphism and mapping of the porcine pyruvate			

[illegible]

Db 4 LQRIQKVLNVRGEIAIRVFRACSELGLKTVAIYSKEDSGSYHRYKADSESYLVGECKPI 63
 Qy 69 KAYLDEIIEIIGAANKVADAIYFGYGLSENALARECAENGITFIPTPEVLDLTGDKS 128
 Db 64 DAYLDEIGIIEIAKSNHVDIAHPGYGLSENIFAKRCEBEGIFIGPKSKHLDMPGDKV 123
 Qy 129 RAVTAAGKAGLPVLAEST-PSKNIDEIVKSAEGOTYPIFKAVAGGGGGRMRVAPDEL 187
 Db 124 KARTOAOALQIPVPGSDGPNVSEVEKFAEKYDPIIIKASIGSGGGRMRVIRISEEL 183
 Qy 188 RKLATEAREAAAFGCGAVVVERAVINPOHIEYQILGDHDTGEVHLYERDCSLQRHQK 247
 Db 184 GESYNRAKSEAKAAGFNDVVEKPKHIEYOILADEEGNVHLYERDCSVQRHQK 243
 Qy 248 VWEIAPAGHLPDELDRICADAVKFCRSIGYCGAGTVEFLVDEKGNHVFLEWNPRTQVEH 307
 Db 244 VWEIAPSVLSDDLQRICDAVXLTKNVNLYNAGTVEFLVKD-DEYFIEVAPRQVEH 302
 Qy 308 TVTBEVTEVDLVKQMRILAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNQRPD 364
 Db 303 TITMITGVDIVQSOILLADGSHLSKMGVGPKEEVVHGFALQSRVTTEDPLNNFMPD 362
 Qy 365 TGTITAYRSPGAGVRLD-GAAQILGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEF 423
 Db 363 TCKIMAYSRSGGGFGVRLDTGNSFGAVITPYDLSLVKVTWALTFFQAAKMERNUKEF 422
 Qy 424 TVSGVATNIGFLRALLREDEFTSKRIATGFIADPHILLQAPPADDEQGRILDYLVADVTN 483
 Db 423 RIRGKTNI PLENVVXKHNFLSGEYDTSFIDASPELFLPKRDRGKTMKLYIGTVTN 482
 Qy 484 KPHGVRPKDVA-PIDKLPNTK-DLPLPRGSRDLKQLGPAAFARDLBEQDALAVTTTF 541
 Db 483 GPPGVGKKEKIPFPDARIPCLKHSEPIQNGTKQILDERGADGLVKVQDQKRVLLTDTTF 542
 Qy 542 RDAHOSLLATRVRSFALPAEAAVAKLTPELLSVEANGATYDVAMRFLFEDPMDRDLDEL 601
 Db 543 RDAHOSLLATRIWDLQIAEPARMLPNLFSAEMWGATFDVAYFLKEDPWERLLDL 602
 Qy 602 REAMPNVIQMLRGNVTGYTPDSCVAFVKEAASGVDIIPRFDALNDVQMRPAI 661
 Db 603 REKNENVLQMLRNSNAVGVKNYPDNLIQFVCSAQAGIDVRIFDLSLNWVEGMRVAI 662
 Qy 662 DAVLENTAVAAVAMAYSGLSDPNEKLYTLDYLUKMAEIVKSGAHILAIKDWAGLLRP 721
 Db 663 DAVRTG-KIABATWCYTGDIHDLPSKDYLDNYKNLAKELASGAHLIGKDWAGLLKP 721
 Qy 722 AAVTKLVTALREDFDLPVHVTHTDTAGGOLATYFAAQAQADADVGASAPLSGTTSPSL 781
 Db 722 NAAVDLVSAKETSIPILHLDHDTSGNGILTYTKRAIEAGVDIVDVAVSSMAGOTSQPSA 781
 Qy 782 SAIVAAFAHTRDITGLSEAVSDLEPYEAVRGLVPESGTPGPTGRVYRHEIPGGOLS 841
 Db 782 NTLIYALGNBERQPDVNIIDSLSEKLSHYWEDVVKYIYAPFESGMNAPHTVEVYHWPGGQYS 841
 Qy 842 NLRAQATALGLADRFELIEDNYAANVEMLRPTKVTTPSSKVVGDIALHLVAGVDPADFA 901
 Db 842 NLQQAQAVGLGDRFDEVKVYRVRVNDMGDVKVTPSSKVVGDMALEPMVQNHILTEQDVL 901
 Qy 902 ADPKVDIPDSVIAFLRGLGNPPGWPPEPLRTRALEGRSEKAPLTVPEEQAHLAD 961
 Db 902 ERGHAMDFGVSVEWEMFSGDLGQPYGGFKELQKIL----KKEPLTVRPELLEPVDFE 957
 Qy 962 DSKERR-NSLNR-----LLFPKPTBEFLHRRRFGNTSALDDREFFYGLVEGRET 1010
 Db 958 ALKEELFHLKREVTTFDVVAVALYPKVPMDEKVAELYGNVSLDTPTFFYGMRLGSEI 1017
 Qy 1011 LRLPLDVRPLVRLDAISEPDDKGNRVNVANVNGQIRMRVRDRSVESVTATAEKADSS 1070
 Db 1018 DVEIEGGKT-LMWKLVSIQEPDQGNRVLYLBFNGQPREIIVKDESVRKATVAQRVKGNRE 1076
 Qy 1071 NKGHVAAFPAG-VVTVTVTAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRVVVVPAATK 1129
 Db 1077 NPNHISATPEGTIVKVVVKEGDEVKGDSSMAITEAMKMETTVQAPFNGKVKVYVNDGDA 1136

Qy 1130 VEGDGLIV 1137
 Db 1137 IQTGDLII 1144
 RESULT 15
 Q81MT6 PRELIMINARY; PRT: 1148 AA.
 AC Q81MT6,
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Pyruvate carboxylase.
 GN PYC OR BA4157.
 OS Bacillus anthracis (strain Ames).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CX NCBI_TaxID=198094;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22608414; PubMed=12721629;
 RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
 RA Nelson K.E., Tetelin H., Fouts D.E., Eisen J.A., Gill S.R.,
 RA Holzapfel E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
 RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
 RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
 RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
 RA Benton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.F.,
 RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
 RA Hazen A., Claine R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
 RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
 RA Fraser C.M.;
 RT "The genome sequence of Bacillus anthracis Ames and comparison to
 RT closely related bacteria.";
 RL Nature 423:81-86(2003).
 DR EMBL; AB017037; AAP27881.1; --
 DR TIGR; BA4157; --
 DR GO; GO:0005737; C:cytoplasm; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0018874; F:lipase activity; IEA.
 DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR005482; Biotin carb C.
 DR InterPro; IPR000089; Biotin lipoyl.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR005481; CPase_L_N.
 DR InterPro; IPR000891; HMG-Like.
 DR InterPro; IPR003379; PYC_OADA.
 DR InterPro; IPR005930; Pyruv. carbox.
 DR Pfam; PF02785; Biotin carb C; 1.
 DR Pfam; PF00364; biotin lipoyl; 1.
 DR Pfam; PF00289; CPase_L chain; 1.
 DR Pfam; PF02786; CPase_L_D2; 1.
 DR Pfam; PF00682; HMG-Like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR TIGRfams; TIGR01235; pyruv. carbox; 1.
 DR PROSITE; PS00866; CPASE_1; 1.
 DR PROSITE; PS00867; CPASE_2; 1.
 KW Pyruvate; Complete proteome.
 SQ SEQUENCE 1148 AA; 128573 MW; 57B97F8D9D1287BF CRC64;

Query Match 43.8%; Score 2526; DB 16; Length 1148;
 Best Local Similarity 47.2%; Pred. No. 7.1e-126;
 Matches 542; Conservative 178; Mismatches 402; Indels 26; Gaps 13;
 Qy 9 LPAPKILVANRGEIAIRVFRACSELGLKTVAIYSKEDSGSYHRYKADSESYLVGECKPI 69
 Db 4 LQRIQKVLNVRGEIAIRVFRACSELGLKTVAIYSKEDSGSYHRYKADSESYLVGECKPI 63
 Qy 69 KAYLDEIIEIIGAANKVADAIYFGYGLSENALARECAENGITFIPTPEVLDLTGDKS 128
 Db 64 DAYLDEIGIIEIAKSNHVDIAHPGYGLSENIFAKRCEBEGIFIGPKSKHLDMPGDKV 123

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:40:22 ; Search time 27 Seconds
(without alignments)
4061.420 Million cell updates/sec

Title: US-10-045-072-2
Perfect score: 5788
Sequence: 1 MSHHSSTLPAPFKILVANR.....RVVPAATKVEGGLIVVVS 1140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues 283366
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: piri: *
2: piri2: *
3: piri3: *
4: piri4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3712.5	64.1	1127	2 D70671	pyruvate carboxyla
2	2541.5	43.9	1178	1 A47255	pyruvate carboxyla
3	2539.5	43.9	1178	2 JC4391	pyruvate carboxyla
4	2525.5	43.6	1150	2 A83978	pyruvate carboxyla
5	2524	43.6	1148	2 F69685	pyruvate carboxyla
6	2517.5	43.5	1178	1 JC2460	pyruvate carboxyla
7	2493	43.1	1146	2 AH1208	pyruvate carboxyla
8	2481.5	43.0	1144	2 D97227	pyruvate carboxyla
9	2480	43.0	1146	2 AC1505	pyruvate carboxyla
10	2488	43.0	1150	2 G89881	pyruvate carboxyla
11	2473.5	42.7	1174	2 AE2911	pyruvate carboxyla
12	2473.5	42.7	1174	2 C97686	pyruvate carboxyla
13	2464.5	42.6	1158	2 AE3285	pyruvate carboxyla
14	2464	42.6	1175	2 T20346	pyruvate carboxyla
15	2457	42.4	1178	1 QYBYP	pyruvate carboxyla
16	2443	42.2	1185	2 T39734	pyruvate carboxyla
17	2436	42.1	1180	2 S46094	pyruvate carboxyla
18	2419	41.8	1195	2 T43735	pyruvate carboxyla
19	2406	41.6	1137	2 E86708	pyruvate carboxyla
20	2115.5	36.5	984	2 T44608	pyruvate carboxyla
21	1079	18.6	501	2 D64453	biotin carboxylase
22	1063.5	18.4	477	2 G70437	biotin carboxylase
23	1036.5	17.9	447	2 A53311	biotin carboxylase
24	1036.5	17.9	447	2 AH1923	biotin carboxylase
25	1035	17.9	472	2 A70432	biotin carboxylase
26	1019	17.6	506	2 D69277	biotin carboxylase
27	997	17.2	491	2 A69123	biotin carboxylase
28	981.5	17.0	448	2 S74380	biotin carboxylase
29	976.5	16.9	471	2 G82966	probable biotin ca

30	975.5	16.9	1095	2 B83471	probable pyruvate
31	971.5	16.8	1078	2 D87647	hypothetical prote
32	961.5	16.6	447	2 B97338	biotin carboxylase
33	961	16.6	677	2 AC2997	hypothetical prote
34	959	16.6	667	2 F98286	hypothetical prote
35	957.5	16.5	444	2 C70444	biotin carboxylase
36	957	16.5	449	2 AI0912	biotin carboxylase
37	951.5	16.4	455	2 B86722	biotin carboxylase
38	949	16.4	449	2 AD0445	biotin carboxylase
39	948.5	16.4	539	2 T07093	acetyl-CoA carboxy
40	946	16.3	444	2 T44813	biotin carboxylase
41	946	16.3	448	1 F64105	biotin carboxylase
42	942.5	16.3	455	2 C95049	acetyl-CoA carboxy
43	942	16.3	450	2 A69581	acetyl-CoA carboxy
44	941.5	16.3	455	2 A97920	biotin carboxylase
45	939.5	16.2	457	2 H71553	probable biotin ca

ALIGNMENTS

RESULT 1
D70671
pyruvate carboxylase (EC 6.4.1.1) - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 11-Jan-2002
A:Accession: D70671; S73055
R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Article: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70671
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1127 <COL>
A:Cross-references: GB:Z83018; GB:AL123456; NID:G3261671; PIDN:CAB05410.1; PID:G169486
A:Experimental source: strain H37RV
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, September 1994
A:Description: Mycobacterium tuberculosis cosmid tbc2.
A:Reference number: S73053
A:Accession: S73055
A:Molecule type: DNA
A:Residues: 1-353, TRAGSARCDPPAVPVSAWTAATWRNQVLRHAGQADLS, 396-1115, 'EMRAETCWNW' <S
A:Cross-references: EMBL:000024; NID:G560506; PIDN:AAA50948.1; PID:G560527
C:Genetics:
A:Gene: pca; pyc
A:Start codon: GTG
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindin
C:Keywords: biotin binding; ligase; mitochondrion
F:4-457/Domain: biotin carboxylase homology <BCH>
F:1055-1127/Domain: lipoyl/biotin-binding homology <LPB>
F:1093/Binding site: biotin (lys) (covalent) #status predicted

Query Match		64.1%	Score 3712.5;	DB 2;	Length 1127;
Best Local Similarity		64.4%	Pred. No. 2.4e-196;		
Matches 730;		Conservative 153;	Mismatches 239;	Indels 11;	Gaps 6;
QY	12	FKKILVANRGEIATVAFRAAL	ETGAATVATVPREDRGSFHSFASAEAVRIGTEGSPVKAY	71	
DB	2	FSKVLVANRGEIATVAFRAAL	YELGVGTVAVPYEDRNSQHRKKADESQIGDIGHPVHAY	61	
QY	72	LDIDRITIGAAKKVKADAIYP	GYGFLSENAQLARECAENGITFPGTPEVLDITGDKRAV	131	
DB	62	LSVDIVATARRAGADAIYP	GVGFLSENPDLAAACAAAGISFVGPFAEVLGAKNSRAI	121	
QY	132	TAAKVAGLPVLAESTPNS	KNDIVKSAEQVPIPVKAVAGGGGGMFVASPDLRKLKLA	191	
DB	122	AAAREAGLEPVLMSAP	SASVDELLSVAAGMPPFLPVKAVAGGGGGMRRVGDIAALPEAI	181	

QY 788 FAHTRDTGLSLVANSLEDFEYWEAVRGVLYLPESGTPGPTGR--VYRHEIPGGQSLNRA 845
 Db 817 TKGTPDTEVPVDFYDSEYWEAGVGLYAAADCTATWKSNGSDYVNEIPGGQYTNLHF 876
 QY 846 QATAGLADRFELIEDNYAANVEMLRPTKVTTPSSKVGDLALHVGAGVDPADPAAPQ 905
 Db 877 QAHSMGLSGKFEVKAYVEANQMLGDLIKVTFSSKIVGDLAQFVWQNGLSRAEAQAQAE 936
 QY 906 KYDIPDSVTAFLRGLNPGGPEPLRTRAL-----EGRSEGKAPLTPVEPEEQALHD 959
 Db 937 EUSFFRSVEFLQGYIGIPHGFPFPRSKVKLDKPRIEGRGASLPPLNLEKELDIL 996
 QY 960 A-DDSKERNLSNRLFPKPTPEEFLEHRRFRFGNTSALDDREFFYGLVGEFRETILRLPDR 1018
 Db 997 RHGEVTPEDVLSAAMYDVFAQKDFATFGLDLSNLTFLQPKIAEPEVELEK 1056
 QY 1019 TPLVRLDAISPPDDKGMNVANVNGQIEPRVRDRSVESVTAETAKDSSNKGHVAAP 1078
 Db 1057 T-LHIXALAVSRLNAGQOVFFELNGQLRSILVKTQAMKEMHFPKALDKVKGQIGAP 1115
 QY 1079 FAG-VVTVTVAEDEVKAGDAVAIIEMKWEATITASVDGKIDRVVVPAAATKVEGGDLIV 1137
 Db 1116 MPCKVIDIKVAGDKVAKGQPLCVLSAMKMTWTWSPMEGTIRKVVHTKDMTLEGDDLIL 1175
 QY 1138 VV 1139
 Db 1176 EI 1177

RESULT 3
 JC4391
 pyruvate carboxylase (EC 6.4.1.1) precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Jan-1996 #sequence revision 25-Apr-1997 #text change 11-Jan-2002
 R:Accession: S68252; S72393; JC4391; S06440
 R:Jitrapakdee, S.; Booker, G.W.; Cassady, A.I.; Wallace, J.C.
 Biochem. J. 316, 631-637, 1996
 A:Title: Cloning, sequencing and expression of rat liver pyruvate carboxylase.
 A:Reference number: S68252; MUID:96257760; PMID:8687410
 A:Accession: S68252
 A:Molecule type: mRNA
 A:Residues: 1-1178 <JITI>
 A:Cross-references: EMBL:U36585; NID:g1040973; PIDN:AAC52668.1; PID:g1040974
 A:Residues: 1-1178 <JITI>
 A:Accession: S72393
 A:Molecule type: Protein
 A:Residues: 489-505 <JIT2>
 A:Experimental source: Liver
 R:Lehn, D.A.; Moran, S.M.; MacDonald, M.J.
 Gene 165, 331-332, 1995
 A:Title: The sequence of the rat pyruvate carboxylase-encoding cDNA.
 A:Reference number: JC4391; MUID:96096548; PMID:8522203
 A:Accession: JC4391
 A:Molecule type: mRNA
 A:Residues: 1-221, 'P', 223-865, 'D', 867-976, 'G', 978-1178 <LEH>
 A:Cross-references: GB:U32314; NID:G929987; PIDN:AAA96256.1; PID:G929988
 A:Experimental source: Liver
 R:Thampy, K.G.; Huang, W.Y.; Makil, S.J.
 Arch. Biochem. Biophys. 266, 270-276, 1988
 A:Title: A rapid purification method for rat liver pyruvate carboxylase and amino acid
 A:Reference number: S06440; MUID:89024676; PMID:3176228
 A:Accession: S06440
 A:Molecule type: protein
 A:Residues: 'SG', 23-25, 'PL', 28-29, 'LL', 32-34, 'P', 1134, 'A', 1136-1137, 1139-1152, 'T', 1154-1155
 C:Comment: This enzyme is located in the mitochondrial matrix and catalyzes the conversion of pyruvate to oxaloacetate, an intermediate that enters the citric acid cycle.
 C:Genetics:
 A:Gene: pc
 A:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
 C:Keywords: biotin binding; gluconeogenesis; homotetramer; ligase; mitochondrion
 F:1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
 F:21-1178/Product: pyruvate carboxylase #status predicted <MAT>
 F:39-494/Domain: biotin carboxylase homology <BCH>

F:1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
 F:1144/Binding site: biotin (lys) (covalent) #status experimental

Query Match 43.9%; Score 2539.5; DB 2; Length 1178;
 Best Local Similarity 46.8%; Pred. No. 9.5e-132; Indels 29; Gaps 12;
 Matches 542; Conservative 185; Mismatches 410; Indels 29; Gaps 12;

QY 1 MSTHTSTSLPA-----FKILVANRGEIAVRAAFRAALETGAATVAIYPEEDRG 48
 Db 14 LGVRRSSTAPVSPNVRRLLEYKPIKVMVANRGEIAIRVFRACTELGIRTVAVYSEQDTG 73
 QY 49 SFRSFASEAVRIGTSGSVKAYLDIDELIIGAACKYKADAIYGVYGLFENALAECAE 108
 Db 74 QMERQKADAYLIGRGLAPVQVYLHPIIDIKVAKENGVDVHGVYGLFESRADFAACQD 133
 QY 109 NGITFIGPTPEVLDTGCKSRVAAAKAGLPVL-AESTPSKNIDIVKSAEGQYPIFV 167
 Db 134 AGVRFIGPSFVVRKMGDKVEARALAIAGVPVPGTNGSPINSLEHAEFSNTYGPPIIF 193
 QY 168 KAVAGGGGRRMRFVSPDELRLKLTAEASRAEAAFGDGAIVYVERAVINQHIEVQLLDGH 227
 Db 194 KAAVGGGGRMRVHVSYELEENYTRAYSEALAFNGALFVEKFEKPHIEVQLIGQ 253
 QY 228 TGEVWHLRYDCSLQRRHQKVVIEIPAQHLDPDLRLDRICADAVKFCRSIGYQGAGVPEFL 287
 Db 254 YGNILHLRYDCSLQRRHQKVVIEIPAQHLDPDLRLDRICADAVKFCRSIGYQGAGVPEFL 313
 QY 288 VDEKGNHVFTEMPRIQVEHTVTEVDLVKQWRLAAGATILKELGLTQDKIKTHGAA 347
 Db 314 VDKGHVYFVANSRLQVEHTVTEITDVLVHAQHVSEGRSLPDLGLRQENIRINGCA 373
 QY 348 LQCRITTEDNNGFRPDTGITAVRSPGAGVRLDGAAL-GGEITAHFDSMLVYKVTCTRG 406
 Db 374 IQCEVTTEDPARSFOPTGTGRIEVRSGEGWIRLDNASAFQGAIVSPHYDLSLVKVIAG 433
 QY 407 SDFETAVARAQALAEPTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPA 466
 Db 434 KDHPATAATKMSRALAEPRVGVKTNIPFLQNLVNNQFLAGIVDTQFIDENPELFLQRLPA 493
 QY 467 DDEQGRILDVLYADVTVNKHGVRPKDVA-APIDKL-PNIKDLPLRGSRDLKQLQAPAA 524
 Db 494 QNRACKLHLVGHVWNGPTTPIPVKVSFVDPVPIVPPVFIQPPAGFRILLAREGEGF 553
 QY 525 ARDLREODALAVDTTTFERDAHQSLATRVRSFALKPAAEAVALKTPBELLSEVAGWATYD 584
 Db 554 ARAVRNHQGLLLMDTTFERDAHQSLATRVRTDHLKTIAPYVAHNFNNLFSIENWGGATFD 613
 QY 585 VAMRFLEDPWDLDELREAMPVNIOMILRGNTVGYTPYDPSVCRAPFVKEAASSGVDI 644
 Db 614 VAMEFLYCEPMRRRLQELRELIPNIPQMLLRGANAVGTYTPDNNVFKFCEVAKENGMDV 673
 QY 645 FRIFDALNDVDSQMRPAIDAVLETTNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKVAEIVK 704
 Db 674 FRIFDSLNLPLNMLLGMEEAGSAG-GVVEAAISYTGVDADPSRTKYSLEYMGLAEELVR 732
 QY 705 SGAILAIKDWAGLLRPAANTKLVTALRRFF-DLPVHVHTDGTAGGOLATVFAAAQAGAD 763
 Db 733 AGTHILCKDWAGLLKPAACACTMLVSSLRDRFPDPLPHIHTDTSVGSVAAMLACAQAGAD 792
 QY 764 AVDGASAPLSGTTSPSLSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGT 823
 Db 793 VDVVAVDSMSGWTQPSMGALVACTKGTPLDTEVPLERVFDYSEYWEAGARLYAAFDCTA 852
 QY 824 PGPTGR--VYRHEIPGGQSLNRAQATLGLADREFELIEDNYAANVEMLRPTKVTTPSSK 881
 Db 853 TMSKNSDVSVEIPIGGQYTNLHFQAHSMGLSGKFEVKAYVEANQMLGDLIKVTFSSK 912
 QY 882 VVGDLALHLVAGVDPADFAADPKYDIPDSVIAFLRGELGNPPGWPPELRTAL---- 937
 Db 913 IVGDLAQWVQNGLSRAEAAQAESELSFPRSVVEFLQYIGIPHGPEPFRSKVLXDL 972
 QY 938 --EGRSEGKAPLTPVEPEEQALHD--DDSKERNLSNRLFPKPTPEEFLEHRRFGNTSA 994

Db 973 RIEGRPGASUPPLNKLKLEKDLIDRHEEVTPELVSAAMYPDVFAQFQKDFATFGPLDS 1032
 Qy 995 LDRREFPYGVEGRETLILPDPVRLPLVLLDAISPPDDKGMNVVNVVNGOIRPMVRD 1054
 Db 1033 LNTLFLQGPXIAEBEVELEGRKT-LHICALAVSDLNPAQORQVFFELNQLSILVKD 1091
 Qy 1055 RSVESVTATAKADSNKHGVAAPFAG-VVTVVAGDEVKAGDAVAIIEMKWEAITA 1113
 Db 1092 TQAMKEMHFPKALKDVKGQIGAMPKGVKIDVYKVAAGAKVWKGQFLCVLSAMKMETVVT 1151
 Qy 1114 SVDGKIDRVVPAATKVEGGDLIVVV 1139
 Db 1152 PMESGTRKVVHTKDMTEGGDDLILEI 1177

RESULT 4
 A:83978
 C:Species: pyruvate carboxylase pycA [imported] - Bacillus halodurans (strain C-125)
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: A83978
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: A83978
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1150 <STO>
 A:Cross-references: GB:AP001516; GB:BA000004; NID:G10175192; PIDN:BA06344.1; GSPDB:GN03
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: pycA
 C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 43.6%; Score 2525.5; DB 2; Length 1150;
 Best Local Similarity 46.8%; Pred. No. 5.4e-131;
 Matches 539; Conservative 185; Mismatches 400; Indels 27; Gaps 13;

Qy 9 LPAPFKILVANRGSIIVAFRAALETGAATVAIYREDRGSPHRSFASFAVRICTEGSPV 68
 Db 4 LKNIKVLVANRGSIIVAFRAALETGAATVAIYREDRGSPHRSFASFAVRICTEGSPV 63
 Qy 69 KAYLIDIDEITGAARKVADAIYGYGFLSENACLAECABENITRIGTPEVLDTGDKS 128
 Db 64 EAVLDIEGIIETIAKRGVDAIHGPGYGLSENIEFAKRCHEGIIIFIGPELHVMFGDKV 123
 Qy 129 RAVTAACKAGLPVLAEST-PSKNIDRIVKSAEQTYPIFKVAVAGGGRGMRFPVSPDEL 187
 Db 124 QAREQAKNLPIVPGSDGVSSLEVDKAFKKGHPFFIKALGGGGRGMRFPVSENDV 183
 Qy 188 RKLATEASREAAFGVAVYERAVINPOHIEVQILGDHTGEVVLHYERDCSLQRHQK 247
 Db 184 QESYERAKSEAKAAGNDEVYKFIENPKHIEVQILADKHGNTLHLYERDCSVQRHQK 243
 Qy 248 VWEAPACHLDPELRICADAVKFCRSIGYCGAGTVFELVDEKGNHVFIEWNPRIQVEH 307
 Db 244 VWEAPSVLSIEDVREICQAAVOLAEVNNVYAGTVFELVDEKGNHVFIEWNPRIQVEH 303
 Qy 308 TVTEVEVDLVKQVRLAAGATL--KEGL-TQDKIKTHGAALOCRTTDPNNGFPDP 364
 Db 304 TITMTVTCIDVOSQFLADGEHLGDRGLGIPKQEEIVCHGYAIGSRVTTEDPSNGFLPD 363
 Qy 365 TGITAYRSPGGAGVRLD-GAALGGEITAHFDSMLVKTGCSGDFETAVARAQALAEF 423
 Db 364 TGRINAVRSGGFGVRLDAGNGEQGAVIPYDLSLVKSTWALLTFEGNAKVLNREF 423
 Qy 424 TVSGVATNIGFLRALREEDFTSKRIATGFIADPHLLQAPPADDEQGRIDYLDVTVN 483
 Db 424 RIRGIKNTIAFLNVVHQRFGLSGEYNTSFIDQTELPFVFPKRKDRGTMLSFGETIVN 483
 Qy 484 KPHGVKPDVAAPIDK--LPNIK-DLPLRGSRDLKQLGPAFAFADRLREQDALAVTTT 540

Db 484 GYFGLG-KTKKPVDFKPPVKLKLSEPIPDGKTQILDHQHPEGLAKWVKSQKHVLLTDTT 542
 Qy 541 PRDAHQSLATRVRSFALKPAAEAVALKTPBELLSEVWAGGATVAVRFFEFDFWDLDE 600
 Db 543 PRDAHQSLATRVTRHDLKQIAETARLLPNLFSWEMWGGATPDVAVRFLHEDFWELL 602
 Qy 601 LREAMPNNIOMLGRNTVGYTYPDSVCAFAFVKEAASSGVDFRIFDALNDVSQMRPA 660
 Db 603 LRKKAPNLFOWLLFRASNAVGYKNYPDNLREFYDKSANAGIDVFRIFDSLNWVEGKLA 662
 Qy 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEIEVKSGAHILAIKDMAGLLR 720
 Db 663 IEAVGEAN-KIAEATICVTGDILOSSRPKYDLAYKKLAKLEAEAGAHILIKDMAGLLK 721
 Qy 721 PAATVKTALAREPDLPHVHTHTDTAGGQLATYFAAQAQADAVDCASAPLSOTTSOPS 780
 Db 722 PFAAYCOLVAELKQVTTIPVHLHTDTSGNGIFTVARAIEAGVDIVDVAVSSMAGLTSQPS 781
 Qy 781 LSAIVAAFAHTRRDGTLSLEAVSDLEPYWEAVRGLYLPFESGTPPGTGRVYRHEIFGGQL 840
 Db 782 ANSLYVALADSERQPNVNITALEQLAEFWETRFYAGFESGKNAPHTVEVHEMPGGQY 841
 Qy 841 SNLRAQATGALGLADREFELIEDNYAAVNMELGRPTKVTSSKVVGDALHLVAGVDPADF 900
 Db 842 SNLQQAQAVGLGHRNNEVKNTVNDMFDGVVKVTPSSKVGDMALVWQNDLTBEV 901
 Qy 901 AADPOKYDIPDSVIAFLRCELGNPPGWPPELRTRALEGRSEKAPLTYPEEEQALHA 960
 Db 902 YENGHKLDPDSVVEFFEGQLGQVQGPQPKLQBIILKGRK---PITNRFGENMEPIQF 957
 Qy 961 DDKSKR-ENSLNR-----LLPKPTEEFLEHRRRTGNTSALDDREFFVGLVGRE 1009
 Db 958 EAIKELYNKLDQVTSHDILSYALYKVFWEFRFQTGDVSVLDTPTFFYGLRPEE 1017
 Qy 1010 TLIRLPDVRTLLVRLDAISPPDDKGMNVVNVVNGOIRPMVRDRSVESVATAEKADS 1069
 Db 1018 IEVEIEQKTLIVKXIFSLSKPQDDGNRIVYFELNGQPREVLIKDQSVKTSIISRPKADK 1076
 Qy 1070 SNKGHVAAFPAG-VVTVVAVSGDEVKAGDAVAIITAMKWEAITATASVDGKIDRVVPAAT 1128
 Db 1077 SNPNHIGSMPTGVVAVKALVEKDKVQKQDHLMITAMKWEITVQAPFDGEVVALVHKGD 1136
 Qy 1129 KVEGGDLIVVV 1139
 Db 1137 AIQTGDLIIIEV 1147

RESULT 5

pyruvate carboxylase (EC 5.4.1.1) pycA [similarity] - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Aug-2000

C:Accession: F69685

C:Authors: Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
 R.; Kunst, F.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 C.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, E.
 Nature 390, 249-256, 1997

A:Authors: Poulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galle
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoisappel, S.; Hosono, S.; Hullo, M.F
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
 A.; Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
 akeuchi, M.; Tanakechi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
 A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Banchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: F69685
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1148 <KUN>

A;Cross-references: GB:Z90111; GB:AL009126; NID:G2633699; PIDN:CAB13359.1; PID:G2633857
A;Experimental source: strain 168
C;Genetics:
A;Gene: pycA
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
F;8-465/Domain: biotin carboxylase homology <BCH>
F;1073-1146/Domain: lipoyl/biotin-binding homology <LPB>
F;1112/Binding site: biotin (lys) (covalent) #status Predicted

Query Match 43.6%; Score 2524; DB 2; Length 1148;
Best Local Similarity 46.0%; Pred. No. 6.5e-131;
Matches 533; Conservative 166; Mismatches 392; Indels 48; Gaps 13;

QY 11 APKILVANRGEIAPRAALETGATVATYIPREDGSHRSFASAVRIGTEGSPVKA 70
DB 5 STQKLVANRGETAIRIFRACTELNIRTVAVSKSDSGSYHRYKADAYLVGEGKKPIDA 64
QY 71 YLDIDIEIIIGAAKVKADATVPGYGLSENAQLARECAENGITFIPTPEVLDLTGDKSRA 130
DB 65 YLDIEGIIIDAKENKVDATHPGYPGLSENIHPARRCEEIGVIFGPKSEHLMDFGDKVKA 124
QY 131 VTAAKAGLPVLAEST-PSKNDIEIVKSAGQYPIFVKAVAGGGGRGMRFPVASFDELK 189
DB 125 REQAEKAGIFVPGSDGPAETLEAVEQFGANGYPIIILKASIGGGGRGMRIVRSSEVKE 184
QY 190 LATEASREAAFGDGAVERAVINPOHIEVOILGDHTGEVVHLYERDCSLQRHQKV 249
DB 185 AYERAKSEAKAAGNDEVVVEKLIENPKHIEVQVIGDKGNVHLFERDCSVQRHQVI 244
QY 250 EIAPAHLDELDRICADAKVFCRSIGYGGAGTVFLVDEKGNHVFIEWPRIQVHTV 309
DB 245 EVAPSVLSPELRDQICEAAVALAKNVINAGTVFVLV-ANNEFYFIEVNPVRQVHTI 303
QY 310 TEVTEVDIVKQMLAAGATL--KELGLTODK-IKTHGAALOCRTITTEDPNNGRPTDG 366
DB 304 TEMITGVDIIVQILVAQSHLSKKVNPPEQXDIITIGVATQSRVTTEDPQDNFDPDG 363
QY 367 TITAYSPGAGVRLD-GAAQLGGEITAHFDSMLVMTCRGSDPFTAVARAQALAEFTV 425
DB 364 KIMAYRSGGFGVRLDTGNSFOGAVITPYDLSLLVLTWALTFEQAAAKVRNLQEFRI 423
QY 426 SGVATNIGFLRALIREDEFTSKIAIGFIADPHLLQAAPPADDEQGRILDYLDVTVN-- 483
DB 424 RGKTNIPFLENVAKHKKFTGQDTSFIDTTTELNFNFKQDKRGTKMLTYIGNVTNGF 483
QY 484 -----KPHGVKPKVAAPIDKLPNIKLPPLPRGSRDLKQLGPAAFADLREQDALAV 536
DB 484 PGIGKKEKPAFDPGLGVKVDVQDP-----ARGTKQILDEKGAEGLANVWKEQKSVLL 536
QY 537 TDTTFRDHQSLIATRVRSFALPAEAVAKLTPELLISVEAWGATYDVAMRPELDPWD 596
DB 537 TDTTFRDHQSLIATRVRSFALPAEAVAKLTPELLISVEAWGATYDVAMRPELDPWD 596
QY 597 RLDELREAMPNVNIQMLLRGNITGVYTPYDSCRAFVKEAASGGVDIFRIFALNDVQ 656
DB 597 RLEDLRKEVPNTLFQMLLRSSNAVGYTNPVDNVIKEPVKQSAQSGIDVFRIFDLSLNWKG 656
QY 657 MRPAIDAVLENTAVAEVAMAYSGLSDPNKELVTLDYILKMAEEIVKSCAHILAKDMA 716
DB 657 MTLAIDAVRDTG-KVAEAAICYTDGILDKRTKYDLAYTSMAKELEAAGAHILGDKMA 715
QY 717 GLLRPAATVLTVALRREFLPVHVHTHTAGGOLATYFAAAQAGADAVDGSAPLSGIT 776
DB 716 GLLRPAATVLTVALRREFLPVHVHTHTAGGOLATYFAAAQAGADAVDGSAPLSGIT 776
QY 777 SQPSLSIAVAFATRTDGTLSLEAVSDLEPYWEAVRGLVLPRESGTPGTGRVYRHEIP 836
DB 776 SQPSASGFYHAMEGNDRPERMNVQGVLLSQYWEVSKYSEFESGKMSPTHTIYHEMP 835
QY 837 GGQI-SNLRAQATAGLADRPFLIEDNVAAYNEMIGRPTKVTTPSKVGVGLLALHLVAGVD 896
DB 836 GGQI-SNLQQQAKGVGLGDRWNEKMYRNVDMFGDIVKVTTPSKVGVGLLALVYQVNLIT 895

QY 897 PADFAADPKQYDIPDSVIAFLRGEIENPPGGWPEPLRTRALRSESGKAPLTVPEE--- 953
DB 896 EKDVYKESLDPDPDSVVLFLKGNICQPHGGPEKLOKLILKGOE-----PIIVRPGELLE 951
QY 954 -----EQAHLDDADDKSKERNLNRLFPKPTBEFLEHRRRFGNTSALDDREF 1001
DB 952 PVSFEALKQEFKEQHNLEISD-----QDAVAYALYPKVFTDYVKTTSYGDISVLDTPTF 1007
QY 1002 YGLVEGRETLIRLPDVRTTLLVRLDAISPPDDKGMNVVANVNGQIRPVKVRDRSVESVT 1061
DB 1008 YGMTLGEIEVEIEROKT-LIVKLSIGEPQDPDATVYVYFELNGQPREVVIKDESISKSSV 1066
QY 1062 ATAEDKSSNKGHVAAPFAGVTVTVAAE-GDEVKAGDAVAITIAMKQVETITIASVDGKID 1120
DB 1067 QERLKADRTNPISHAAAMPQGTIVKLAEGATKYNKGDHLMINEAMKQVETTVQAPFSGTIK 1126
QY 1121 RVVVPAAKTVEGGDLVVV 1139
DB 1127 QVHVXNGEPIQTGDLLEI 1145

RESULT 6

JC2460
Pyruvate carboxylase (EC 6.4.1.1) precursor - human
N;Alternate names: pyruvate:carbon dioxide ligase (ADP-forming)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence, revision 18-Aug-2000 #text_change 01-Feb-2002
C;Accession: G01933; JC2460; B27883; S01469
R;Walker, M.E.; Jitrapakdee, S.; Val, D.L.; Wallace, J.C.
submitted to the EMBL Data Library, July 1995
A;Reference number: H00708
A;Accession: G01933
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1178 <WAL>
B;Cross-references: EMBL:U0891; NID:G1101028; PIDN:AAA82937.1; PID:G1101029
R;MacKay, N.; Rigat, B.; Douglas, C.; Chen, H.S.; Robinson, B.H.
Biochem. Biophys. Res. Commun. 202, 1009-1014, 1994
A;Title: cDNA cloning of human kidney pyruvate carboxylase.
A;Reference number: JC2460; MUID:94324922; PMID:8048912
A;Accession: JC2460
A;Molecule type: mRNA
A;Residues: 1-224, 'WP', 227-351, 'A', 353-384, 'PT', 387-485, 'DV', 488-637, 'R', 639-728, 'A', 7
A;Cross-references: GB:S7370; NID:G632807; PIDN:AA31500.1; PID:G632808
R;Lamhonwah, A.M.; Quan, F.; Gravel, R.A.
Arch. Biochem. Biophys. 254, 631-636, 1987
A;Title: Sequence homology around the biotin-binding site of human propionyl-CoA carbo
A;Reference number: A27883; MUID:87212051; PMID:3555348
A;Accession: B27883
A;Molecule type: mRNA
A;Residues: 1083-1178 <LAM>
A;Cross-references: GB:M26122; NID:G189657; PIDN:AAA36423.1; PID:G387003
R;Freytag, S.O.; Collier, K.J.
J. Biol. Chem. 259, 12831-12837, 1984
A;Title: Molecular cloning of a cDNA for human pyruvate carboxylase. Structural relati
A;Reference number: S01469; MUID:85030380; PMID:6548474
A;Accession: S01469
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1135-1178 <FRE>
A;Cross-references: EMBL:K02282; NID:G189655; PIDN:AAA50033.1; PID:G189656
C;Genetics:
A;Gene: GDB:PC
A;Cross-references: GDB:119472; OMIM:266150
A;Map position: 11q11-11q13.1
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: biotin binding; gluconeogenesis; ligase; mitochondrion
F;1-20/Domain: transit peptide (mitochondrion) #status Predicted <TNP>
F;21-1178/Product: pyruvate carboxylase #status Predicted <MAP>
F;39-494/Domain: biotin carboxylase homology <BCH>
F;1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F;1144/Binding site: biotin (lys) (covalent) #status Predicted

Query Match	43.5%	Score 2517.5	DB 1	Length 1178
Best Local Similarity	46.2%	Pred. No. 1.5e-130		
Matches	536	Conservative 194	Mismatches 412	Indels 29
QY	6	STSLPA-----FKKILVANGIIVAFRAALLETGAATVAIIPRDRGSFPHRS	53	
DB	19	TSTAPAAASNVRLEYPKIKKVMANGBIARVFRACTELGIKRTVAIYSEQDTQMGRQ	78	
QY	54	FASEAVRIGTEGSPVKAYIDIOBIIIGAAKKVKADAIYPGYFLSENAQIARCAENGITF	113	
DB	79	KADEAVLIGRLAPQAYLHPDIIKVKAKENNVDVAVHPCGYFLSERADPAQACQDAGVRF	138	
QY	114	IGTPPEVLDTGDKSRAVTAAKAGLPLVLAESTPSKNIDEIVKSAGQTYPIFVKAAG	172	
DB	139	IGSPPEVVRKMGDKVBARAIAAAGYVVPVPGTDAPITSLHEAHFENFTYGFPIIFKAAAY	198	
QY	173	GGGRGMRFVASPDELKRLATEASRREAAAFDGAUVVERAVINPQIEVQIILGDHTGEVV	232	
DB	199	GGGRGNVHVSYEELBENYTRVSEALAAFGNALFVEKFIKPRHIEVQIILGDQYGNIL	258	
QY	233	HLYERDCSIQRHQKVVETAPAOHLDEPLRDRICADAVKFCRIGYCGAGTVEFLVDEKG	292	
DB	259	HLYERDCSIQRHQKVVETAPAOHLDPQRLTSLDSVKLAKOVYENAGTVEFLVDRHG	318	
QY	293	NHVFIEWNPRIQVEHTVTEETVEDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRI	352	
DB	319	KHYFIEVNSLQVEHTVTEETVDVLVHAQIHVSEGRSUPDLGQRKQENIRINGCAIQCRV	378	
QY	353	TTEDPNNGRPDPTGTTIAYRSPGGAGVRLDGAAQL-GEETAHFDSMLVMTCRGSDFFET	411	
DB	379	TTEDPARSPQDTCRIEYFVRSGEMGIRLDSNAFOGAVISPHYDSLVLVKVIAHGKHPT	438	
QY	412	AVARAORALAEFTVSGVATNIGFLRLLREEDTSKRIATGFTADSPHLLQAPPADDEQG	471	
DB	439	AATKMSRALAEFRVGVKTNIAIFLQVNNQQFLAGTVDTQFIDENPELFQLEPAQNRQA	498	
QY	472	RILDYLDATVYNKPHGVRP-KOVAAPIDK-LPNIKDLPGRGSRDLKQUGPAAPARDUR	529	
DB	499	KLLHLYGHVMVNGFTTPIPVKASPSPTDPVPVPAVPGPPAGPAGFDILLRGEGPGEFARVR	558	
QY	530	EQDALAVDTTFRDAHOSLLATVRSFALPAPAAEAVAKLTPELLSVEAMGATVDVAMRF	589	
DB	559	NHFGLLMDTTFRDAHOSLLATVVRTHDLKKIAPYVAHNFSKLFSMENWGAIFDVAMRF	618	
QY	590	LFEDPWDRDLBELREAMPNNVQMLLRGNVTGYTPYPPDSVCRAPVKEAASGVVDIPIRIFD	649	
DB	619	LYFCPWRLQELARELIENIPQMLLRKANAVGVTNPDNVVFKPCEVAKENGMDVFRVFD	678	
QY	650	ALNDVSQMRPAIDAVLETTNVAEAVAYSGDLSDPNEKLYLTLDYLLKMAEEIVKSGAHI	709	
DB	679	SLNYLPNNLQMGAAAGSAG-GVVEAALSYTGVDVADPSRTKYSLOYVGLABELVRAGTHI	737	
QY	710	LAIKDMAGLLRPAATVKLVTALRRFP-DLPVHVHTHTAGQLATYPAQAQADAVDGA	768	
DB	738	LCIKDMAGLLKPTACTVLSLSRRFPDLPLRHTHTDTSAGVAAMLACACAGADVVDVA	797	
QY	769	SAPLSGTTSPSLSAIVAAFAHTRRDPTGLSLEAVSDLEPYWEAVRGLYLPFESCTPGTG	828	
DB	798	ADSNMGSTSPSMLGALVACTRGTPLOTEVPMERVFDYSEYWEAGRGYLAADFCTATWKSQ	857	
QY	829	R--VYRHEIPGGQLSNRAQATALGLADRFELIBDNYYAANVLMGLRPTKTPPSKVVGDL	886	
DB	858	NSDVYENIEPGQVYTNLHFQAHSGMLGSKFKFVKKAYVEANQMLGDLIKVTPSSKIVGDL	917	
QY	887	ALHLVGVDPAADPAADQKVDYDPSVIATFLRGELGNPPGSGWPEPLTRAL-----EGR	940	
DB	918	AQFVQNGLSRAEABAQAELSPFRSVVEFLQGIYGVPHGDFDPFRFSKVLKDLPRVEGR	977	
QY	941	SEKXAPITEYPEEEQAHLDA-DDSKERRNSLNRLLFPKPTTEEFLEHRRRFGNTSALDDRE	999	
DB	978	PGASLPILDLQAFKELVDRHGBVTPEDVLSAAMYPDVFAHPKDFDTATFGLDLSLTRL	1037	

Qy 1000 FFYGLVEGRTLLRLSDVTRTPLLVRDLALISEPDDKGNRVNVVANNVQIQIPMRVDRSVES 1059
Db 1038 FLOQPKIAEFVELERGKT-LHIKALVSDLNRAQGRQVFFELNQLRSILVKDTQAMK 1096
Qy 1060 VTATAEKADSSNGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEMAKMEATITASVDGK 1118
Db 1097 EMHFPKALKDKVKGQICAPMPGKVIDIKVAGAKVAKGQPLCVLSAMKETVVTSPMEGT 1156
Qy 1119 IDRVVVPAATKVGGBLIVVV 1139
Db 1157 VRKHVTKMDLEGDDLILEI 1177

RESULT 7
AH1208
pyruvate carboxylase homolog pycA [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AH1208
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluer, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A.:Title: Comparative Genomics of Listeria species.
A:Reference number: AB1077; MUID: 21537279; PMID: 11679669
A:Accession: AH1208
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1146 <GLA>
A:Cross-references: DB:NC_003210; PIDN: CAC99150.1; PID: g16410474; GSPDB: GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: pycA
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 43.1%; Score 2493; DB 2; Length 1146;
Best Local Similarity 45.8%; Pred. No. 3.3e-129;
Matches 526; Conservative 185; Mismatches 408; Indels 30; Gaps 11;

Qy 13 KKILVNRGEIAVRAFAALETGAATVAIYPREDRGSFHRFSFASEAVRICTEGSPVKAYL 72
Db 5 KKVLVNRGEIAIRVMRACTELKIKTVIYSQEDTGSFHYKSDAEYLVGAGKKPIDAYL 64
Qy 73 DIDSIIGNAKVKADAIYGVGFSENAQLABCAENGIITFIGTPEVLDTGDKSAVIT 132
Db 65 DIENIETAKESGADAIHPGVGFSENIETFAARCEQGGIIFVGPKSXHLDMFGDKIKAKE 124
Qy 133 AAKZAGLPVLAEST-PSKNIDIEIVKSAEGQTYPIFKAVAGGGGRGMRFPVASPDDELKLA 191
Db 125 QALLADIPVPGSGNPVAGIKEVEEFCEKNGYPLMIKASLGGGGRGMRVVESEKHWKESF 184
Qy 192 TEASREBAAGDGAIVVERAVINPQIHIEVQILDGHTGEVWHLYERDCSLQRHKQVEI 251
Db 185 FRASSEAKAAGFNDEVVYEKFCWMPKHIEVQILDGHTGNIIVHLPERDCSIQRHKQVVEV 244
Qy 252 APAQHLDELDRICADAVKFCRIGYQCGAGTVBFVLDEKGNHVFIEENPRIOVHTVTE 311
Db 245 APCNAITSELENRICDAVLMKNVDYINAGTVFELV-EGDDFFYFIEVNPVQVEHTEITE 303
Qy 312 EVTEVDLVKAQMRLAAGATLTKELGLT---QDKIKTHGAALQCRIITDPNNGFPREDGTI 368
Db 304 MITGIDIVQSOLFADGVALHDAQVAIFPKQBDIHGSAIQSIRITTEDPLANNFMDPTGRV 363
Qy 369 TAYRSPGAGAVRLD-GAAQLGGEITAHFDSMLVMTCRGSDFTAVARAQALAEFTVSG 427
Db 364 DTYRSTGGFVRLDAGNGFGQGVTVTPFPVDSLLVXLTKWMTFEQATKRRNLLIEFRIRG 423
Qy 428 VATNTGFURLLRBEDFTSKKIATGFIADHPHLLQAPPADDEQGRDILDLADVTNPKPHG 487
Db 424 VKTNPFFLVNVHPDFDASGNYNTSFIDTTELEFKFPHIRDRGKTGLRYIGNVTVNGPFG 483

QY 1071 NKCHVAAPFAG-VVTVTVAGDEVKAGDAVALIEMKMEATITASVDGKIDRVVVPAAATK 1129
Db 1074 NKKEIGAS:PGNVVKVFKVPGDKVKXGDSLWVIEAMKMETNVSESDGTGGIFVYKGDQ 1133
QY 1130 VEGGDLIV 1137
Db 1134 VQSQQLLV 1141

RESULT 9
AC1565
pyruvate carboxylase homolog pycA [imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1565
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.; Jones, L.M.; Karst, U.
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1565
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1146 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96291.1; PID:gi6413519; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: pycA
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 43.0%; Score 2490; DB 2; Length 1146;
Best Local Similarity 46.0%; Pred. No. 4.8e-129;
Matches 528; Conservative 192; Mismatches 409; Indels 30; Gaps 12;

QY 13 KKLIVANRGEIIVRAFALETCATVATVPREDRGSHRSPASAVRIGTEGSPVKAYL 72
Db 5 KKLIVANRGEIIVRAFALETCATVATVPREDRGSHRSPASAVRIGTEGSPVKAYL 64

QY 73 DIDEITGAARKVADAIYFGYGLSENALAECEANGITFGTPEVLDLTGDKSRVIT 132
Db 65 DIENIEIAKESGADAIHFGYGLSENIEFAKCEQEGIIIVGPKSKHLMDFGDKIKAKE 124

QY 133 AAKAGLPVLAEST-PSKNIDIVKSAEGQTPIFKVAVAGGGRMGFPVSPDELKLA 191
Db 125 QALLADIPVPGSDGVAGIKEVEFEKNGYGLMIKASLGGGGRMGFWESKEHVKESEF 184

QY 192 TEASREAAAFGAGVYVERAVINPOHIEVQILGDTGTVHLYERDCSLORRHOKVVEI 251
Db 185 ERASSEAKAAGFNDVVEKCVNPKHIEVQILGDTGTVHLYERDCSLORRHOKVVEI 244

QY 252 APAQHLDPELRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEWNPRIQVEHTVTE 311
Db 245 APCNAITSLNRNICDAAVKLMKNVDYINAGTVEFLV-EGDNFYFIEVNPVQVSEHTTE 303

QY 312 EVTEVDLVKAQMLAAGATLKELGLT---ODKIKTHGAALQCRITTEDPNNGFRDGTGI 368
Db 304 MITGDIIVSQFIADGYALHQLVAIPKQEDIHGSAIQSRITTEPLNFMFDTGRV 363

QY 369 TAYRSGGAGVRLD-GAAOLGGEITAHFDSMLVMTCRGSDPETAVARAQALAEFTVSG 427
Db 364 DTYRSTGGFVRLDAGNGFQGTWTFPYDSLLVVKLCTWGTMTPEQATRKMRNLIEFRIG 423

QY 428 VATNIGFLRALLEEDFTSKRIATGFIADHPLLQAPPADDSQGRILDYLDADVTNKEHG 487
Db 424 VKTNIPELVNVRHPPAFAGNNTNSFIDTTPPELFPFPHRDRGTGKILRYGVNVTNGPFG 483

QY 488 VRPKDVAAPIDKLPNIKDLP----LPRGRDRKLKQLGPAFAARDLREQDALAVTDTFRD 543
Db 484 IKHRD--KPYVABPRLPKIPYGSQIAPGTQKILDAKPGGVVDWVGKQKEVLLDTTLRD 541

QY 544 AHQSLLATRVRSFALKPAEAENAKUTPELLSVBANGGATYVAMRFLFDFPDWDRIDELRE 603
Db 542 AHQSLLATRVRSKDIPOIADAMAHLPLNNFSEPMGGATFDVAYFLNEDFPVRLTLETK 601

QY 604 AMPNVNIQMLLRGNTVGYTPYDPSYCRAFVKEAASGVVDIFRIFDALNDVQMPAIDA 663
Db 602 QIENVWFQMLLRGANAVGYKNYVDNVIREFVQSQSGVDVFRVFDLSLWIKGMEVSDA 661

QY 664 VLENTAVAEVAMAYSGDLSDPNEKYLTYDYLYLKWABEIVKSGAHILAIKDWAGLLRPAA 723
Db 662 VREAG-KVVEATICYTGDIDDDTRTKYITDYDKMAKELVAQQTILG:KDWAGLLKPOA 720

QY 724 VTKLVTALARREPLPVHVHTHTAGQLATYFAAAQAGADAVDGASAPLSGTTSPSLA 783
Db 721 AYELIGELKXTVDVPFHLHTHTDTSNGIYTYAAASAGVDIVDVASSAMSGATSQPSMTG 780

QY 784 IVAFAHTRDRTGLSLEAVSDLEPYWEAVRGYLPFESCTPGPTGRVYRHEIPGQLSML 843
Db 781 LTYGLVNGRQTNLDNAQNSQIINHVEDVRHYKDFDNALNSPQTEVITHENPGQYTNL 840

QY 844 RAQATALGLADREFELLEDYAAVNEMLGRPTKVTPSSKVVGLALHLVAGVDPAFAAD 903
Db 841 QQAIAVGLGDRWDEVKENVTVVQNFQGDIVKVTSSKVVGLALFMVQNELTEDEVYK 900

QY 904 PQYDIPDSVIAFLRGELNPPGGWPEPLRTRALGRSEKAPLTEVP-----BEE 954
Db 901 GDTIDFPDSVIEPFMGEGIQPGYGGPEKQLVKLGRT---PLADRPCALMEPVNFAEV 956

QY 955 QAHLDAADDSKE--RNSLNRLFPKTEBEFLHRRFRFGNTSALDDREFFYGVGEGETLI 1012
Db 957 KAEKEMGYESEKDVISYILYKVFVDYQEMISKYGDVTVLDTPTFYKGRIGETIV 1016

QY 1013 RLPDVRTPLLRLDAISEPDDKGMNVANVANGQIRPMVRDRSVESVTATAEKADSSNK 1072
Db 1017 ELEKGI-LLIKLSIGEPIDATGTRVYFELNGQPREINIQDMNVQSTVIARRKIDTTNP 1075

QY 1073 GHVAAFPAG-VVTVTVAGDEVKAGDAVALIEMKMEATITASVDGKIDRVVVPAAATK 1131
Db 1076 EHVGNATVGSVIGVVVKGDSVKKGDPLLITAMKMETTIQAPFDEGVSSIVSDGDTIE 1135

QY 1132 GGLIVVVS 1140
Db 1136 SGDLLEIVN 1144

RESULT 10
G89881
pyruvate carboxylase [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89881
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu, ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G89881
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1150 <KUR>
A:Cross-references: GB:BA000018; PID:gl3700915; PIDN:BA842211.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: pycA
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 43.0%; Score 2488; DB 2; Length 1150;
Best Local Similarity 45.2%; Pred. No. 6.3e-129;
Matches 527; Conservative 192; Mismatches 384; Indels 62; Gaps 14;

QY 13 KKLIVANRGEIIVRAFALETCATVATVPREDRGSHRSPASAVRIGTEGSPVKAYL 72

Db 5 KLLLVANRGEIAIRIPRAAAELDISTVAIYSNEDKSLHRYKADBSYLVGSLGPAESYL 64
Qy 73 DIDEIIGAAGKADAIYGYGFLSENAOLARECAENGITFPTPEVDLDTGDKSRVIT 132
Db 65 NIERIIDVAKQANDAIHPGYGFLSENEGAFRCACAEKGFIPGHELDHDMFDKVKART 124
Qy 133 AAKKAGLPVL-AESTPSKNIDEIVKGAAGQTFYFVKAVAGGGGGRMRFVAPDELRKLA 191
Db 125 TAIKADLPVPTGDTGPIKSYELAKEPAEAGFPFMKATSGGGGKGRIVRESELEDAF 184
Qy 192 TEASREAAFGDGAAYVVERAVINPOHIEVOILGDHTGWHLYYERDCSLQRHOKVVEI 251
Db 185 HRAKSEAKGFGNSEVITYINDPKHIEVQVIGDEHGNIVHLFERDCSVQRHOKVVEV 244
Qy 252 APAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLV--DEKGNHVFIEMPRIOVEHPTV 309
Db 245 APSVGLSPTLRQICDAAIQMLMENIKYVAGTVEFLVSGDE---PFEIENVPRVQVEHTI 301
Qy 310 TEZVTEVDLVKQVRLAAGATL--KELGLTQDK-IXTHGAALOCRIITTEDPNNGRPDGT 366
Db 302 TEMVTGDIIVKTLQVLAAGADLFGEEINMPQKDIITLGVAIQCRITTEDPNDFPDGT 361
Qy 367 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVMTCRGSDFETAVARAQALAEFTV 425
Db 362 TIIAYRSSGGFGVRLDAGDGFQGAIEISPYDLSLVKLSHAISFKQAEKQVRSUREMI 421
Qy 426 SGVATNIGFIRALLREEDFTSKRIATGFIADHPHILLOAPPADDEQGRILDLADVTNKP 485
Db 422 RGVKTNIFPLINWKNKFTSGDYTTKFIBETPELFDIQPSLDRGTLEYIGNVTIN-- 479
Qy 486 HGVKPKDVAAPIDKLPNIKDLPLP-----RGSRLKQLGPAAPAFAD 527
Db 480 -----GFPNVEKRPDYELASIPVTSSSKIASFSGTKQLLDEVGPKGVAEM 526
Qy 528 LREQDALAVTDTFRDHAQSLATRVRSFALKPAEAAVAKLTPELLSVEANGATVDVAM 587
Db 527 VKQDDVLLTDTFRDHAQSLATRVATKDMINIAKTADVDFKGFSLNMGWGGATPDVAY 586
Qy 588 RELFEDPDWRLDELREAPNVNQMILRGNTVGYTPYDPSVCRAFFVKEAASGVDFRI 647
Db 587 NFLKENPWERLERKAIIPNVLFQMLLRASNAVGYKYPDNVTHKFOESAKAGIDVRI 646
Qy 648 FDLNDSVQMRPAIDAVLENTNVAEVAAYSGDSDP-NEKLYTDIYDKMAEBEIVKSG 706
Db 647 FDSLNNVDQMKVANEAVQEAQ-KISEGTICYTGDILNPERSNITYTLEYVYKLAKELEREG 705
Qy 707 AHILAIKDMAGLPPAAVTKLVATLRREFDLPVHVHTHTDAGGOLATYPAAAGADAVD 766
Db 706 PHILAIKDMAGLKPAAVELIGELKSAVDLPHLHTHDSGNGLITYKQAIQAGVDIID 765
Qy 767 GASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGP 826
Db 766 TAVASMSGLTSQPSANSIYALNGFPRLRDTIEGMESLSHYMSTVRYTYSDFESDIKSP 825
Qy 827 TGRVYRHEIPGQSLNRAQATAGLADRFELIEDNYAAVNEMLGRETKVTPSSKVVGD 886
Db 826 NTEIYQHEMPGQGSNLSQKSLUGLERFDEVDKMYRNVNIFGDIIVKTPSSKVVGDW 885
Qy 887 ALHLVAGVDPADFAADPKQYDIPDSVIAFLRGLNPGPGWPPEPLTRALEGRSEKAP 946
Db 886 ALYVQNDLDEQSVITDGYKLDPESSVSPFKGIBGQPVNGFNKDLQAVILKQOE---A 941
Qy 947 LTEVEBEEQHLDDADDSKE-----RNSLNRLIIPKTEEBLEHRRFRGTSAL 995
Db 942 LTARPGVELEPDEKVELLEBEOQGPVTEQDIISYLVPKVYEQIQRNQ'GNLSLL 1001
Qy 996 DDRREFFVGLVBGRETLIRLPVTRTLLVRLDAISEPDDKGMNVNANVNGQIRPMVRDR 1055
Db 1002 DTPITFFGMRNGETVEIEI-DKGKRLIKLEITISEPDENGNRITYYAMNGQAAIRYIKDE 1060
Qy 1056 SVESVATAEKADSSNKHVAAPAGVVT-VTVAEQVEKADGDAVAILIEMKWEAITAS 1114
Db 1061 NVHTNANVYKADKSNPSHIGQMPGVSIVTEVKVSVGETVXANQPELLITEAKMKEITIQAP 1120

Qy 1115 VDGKIDRVVVVPAATKVEGGDLIVV 1139
Db 1121 FDGVIKQVTVNNGDTIATGDLII 1145
RESULT 11
AE2911
pyruvate carboxylase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C.Species: Agrobacterium tumefaciens
C.Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C.Accession: AE2911
R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, E.; Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCle Science 294, 2317-2323, 2001
A.Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm ster, E.W.
A.Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A.Reference number: AB2577; MUID:21608550; PMID:11743193
A.Accession: AE2911
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1174 <KUR>
A.Cross-references: GB:AB008688; FIDN:AAL43707.1; PID:g17741236; GSPDB:GN00186
A.Experimental source: strain C58 (Dupont)
C.Genetics:
A.Gene: PyCA
A.Map position: circular chromosome
C.Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
Query Match 42.7%; Score 2473.5; DB 2; Length 1174;
Best Local Similarity 46.6%; Pred. No. 4.1e-128;
Matches 541; Conservative 173; Mismatches 406; Indels 41; Gaps 19;
Qy 8 TLPAPFKILVANRGEIAIRVRAALETGAATVAIVIPREDGSGFHSFASAEVRI 62
Db 20 TVLKISKILVANRGEIAIRVRAALETGAATVAIVIPREDGSGFHSFASAEVRI 79
Qy 63 -TEGSPVKAYLIDIEIIIAKAKVADAIYPCYGLSENAQLARCAENGIFFIPTPEVL 121
Db 80 AKDMGPIESYLSIEVIVAKLSGADALHPGCGLLSESPFVEACNAGITFIPTPTM 139
Qy 122 DLTKGKSAVTAAKKAGLPVLAESTP-SKNTDEIVKGAAGQTFYFVKAVAGGGGGRMRF 180
Db 140 ROLGNKVAARNLAISVDVVPVVPATNPLPDDIAEYERVAEEIGYPVMLKASWGGGGRMRA 199
Qy 181 VASPELREKLATEASREAAAFGDAVYVERAVINPOHIEVOILGDHTGWHLYYERDCS 240
Db 200 IRKEDLAREVTEAKREAAAFGKDEVLEKLEVERAHVESQILGDTHGNVHLEFDCS 259
Qy 241 IQRHOKVVEIAPAOHLDPEDRICADAVKFCRSIGYQAGTVEFLVD-EKGNHVFIEM 299
Db 260 IQRNOKVVERAPAPYLSAQORQELAAVSLKIAAATNYIGAGTVEYLDADTKGFYFIEV 319
Qy 300 NPRIQVHTVTEETVEVDLVKQVRLAAGATL--KELGL-TODKIKTHGAALOCRIITTED 356
Db 320 NPRIQVHTVTEVTGIDIVKQVRLAAGATLGAAGTAEASGVKQEDIRLNGHALOCRIITTED 379
Qy 357 PNNGFRPDTGITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVMTCRGSDFETAVAR 415
Db 380 PEHNFIPDYGRITAYRSASGFIKLDGGTSYTGAVITRYDPLLVKVTAWAPEPEDEAISR 439
Qy 416 AQALAEFTSVSAVNICFLRALLREEDFTSKRIATGFIADHPHILLOAPPADDEQGRIL 475
Db 440 MDRAUREPRIGVATNLTFLEAIGHDSFRNTYITRITIDSTPELFAQVKQDRATKLLT 499
Qy 476 YLADVTVN---KPHG-VRPKDVA--PIDKLPNIKDLPLPRGSRDLKOLGPAAPAFADURE 530
Db 500 YLADVTVNGHPETKGRAXPADKAAKPI--VPYI-DAPTPDGTQQLDLKLGPKGFADWVRN 556
Qy 531 QDALAVTITTFEDAHQSLATRVRSFALKPAEAAVAKLTPELLSVEANGATVDVAMRFL 590

557 EKKVLVDTTMRDGHQSLLATRVTRTHDIARVASVSKALPOLLSECGWGTDFVSMRFL 616
591 FEDPMDRLDELREAMNNVNIOMLLRGRNTVGVTPVDSVCRAFKVKEAASSGVDIFRIFDA 650
617 TEDPHERLSLREGAPNLLQMLLGGANGVGNVNDVVKYFVQAAAGGVDLFRVFD 676
651 LNDVSMRPAIDAVLETTNATAVAEMAYSGBLSDPNEKLYTLDYLLKMAEIVKSGAHIL 710
677 LNWVENRVSMDAIAEEN-KLCEATICTYTGOLLNSAREPKYDKYTNLAVLEKAGAHII 735
711 AJKDMAGLLRPAATVKLVTALRRPDLFVHVHTHTDAGGQLATVFAAQAAGADAVDGASA 770
736 AVKDMAGLLKPAAKVLPKALREATGLFHFHTHTDTSIGISAATVLAADVADDAAMD 795
771 PLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRV 830
796 AFSGNTSQPCLSGIVEALSSESDTGLDTEWIRRIISFYWEAVRQYAAAFESDLKGPASEV 855
831 YRHEIPGGQSLNRAQATLGLADRELIEDNYAANEMLRGPTKVTPESSKVGDLALHL 890
856 YLHEMPGGQFTNLKQASLGLSEHVEVAQYADANRMFGDIVKVTPESSKVGDMALMM 915
891 VGAGVDPADPAADPOKYDIPDSVIAFLRGLGNPFGWPEPLRTRALRSEKAPLITEV 950
916 VSQDLTVADVENPREVSPDSVSMKGLGQSPGWPEALQKAL----KGEKPYTVR 971
951 PEE--EQAHLDADDSDKERRNSLNR-----LLFPKTEEFLEHRRRFGNTSALDDR 998
972 PGSLEDAALDA-ERKVIETKLERKVDDEPFASYLMFKVTFDFALTAETVGPVSLPTH 1030
999 EFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMNRNVANVNGQIRPMVRDRS-V 1057
1031 AVFYGMEDGEELFADIERGKTLVIVN-QASSGIDDKGMVTVFEINGQPRRIKVPDRAHG 1089
1058 EVTATAEAKDSSNKGHVAAFPAGVVT-VTVAGDEVKAGDAVAIIEAMKWEATITASVD 1116
1090 ASGSAVRKAEPGNASHIGAPMGVLSRVFINQGEVKGAGDVLLSIEAMKWEATITASVD 1149
1117 GKIDRVVPAATKVEGGDLIV 1137
1150 GKIAEVLVKGQDIDAKDLII 1170

RESULT 12
C97686
pyruvate carboxylase (U51439) [imported] - Agrobacterium tumefaciens (strain C58, Cereon
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: C97686
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: C97686
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1174 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK88444.1; PID:g15157941; GSPDB:GN00169
C;Genetics:
A;Gene: AGR C 4940
A;Map position: circular chromosome
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 42.7%; Score 2473.5; DB 2; Length 1174;
Best Local Similarity 46.8%; Pred. No. 4.1e-128;
Matches 541; Conservative 173; Mismatches 406; Indels 41; Gaps 19;
8 TLPAFKILVANGETAVAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIG----- 62
20 TVLKISKILVANSEAIRVFRANELGIKTVAIWEEDKLSLHRFKADESYQVGRGPHL 79
63 -TEGSPVKAYLDEIIGAANKVKADAIYPGYGFLSENAQLARECAENCGITFIGTPEVL 121

80 AKDMGPISYSIEEIVIRVAKLSGADAHGPGVGLLSESPFVEACNKAGIIFIGTPTDM 139
122 DLTGOKSRVATAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTPYIFVKAVAGGGGRWRF 180
140 ROLGNKVAARNLAISVDVPPVPATNPLPDDIAEVRMAEEIGYPVWLKASWGGGGRGMA 199
181 VASPELRLKLAIEAREAAAFDGAIVYVERAVINPQHIEVOILGDHGTGEVHLYERDCS 240
200 IRKEDLAREVTEAKREAAAFKDEVLKELVERARHVESQILGDTHGNVHLLFERCS 259
241 LORRHQKVVEIAPAQHLDPEDLRICADAVKPCRSIGYQAGCTVEFLVD-EKGNHVFTEM 299
260 IQERNKVVERAPAPVLSAQOQELAAVSLKIAAATNVIAGCTVEYLMADATGKFFIEV 319
300 NPRIQVEHTVTEVTEVDLVKQMLAAGATL--KELGL-TODKIKTHGAALQCCITTED 356
320 NPRIQVEHTVTEVTEVDLVKQMLAAGATL--KELGL-TODKIKTHGAALQCCITTED 379
357 PNNRFPEDTGTITAVRSPGAGVRLD-GAAQGLGGBITAFDFDSMLVKMTCRGSDFETAVAR 415
380 PEHNFIPDYGRITAYESASGFGIRLDGGTSYGTAVITRYIDPELLVKVTAWEPEDEAISR 439
416 AQALAEFTVSGVATNIGLIRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILD 475
440 MDRALREPRIRGVATNLTFLEAIIGHDSFRNNTYTRFIDSTPELFAQVKRQDRATKLLT 499
476 YLADVTVN--KPHG-VRPKDVAAPIDKLPNIKOLPLPRGSRDRKLQGLPAAFAARDLRE 530
500 YLADVTNNGHPETKGRAPADKAKPI--VPYI-DAPTEGDKQLLDLXGPGFADWMEN 556
531 QDALAVDTTFRDAHOSLLATRVSRFALKPAEAAVAKLTPELLSVEMAGGATYIVAMREL 590
557 EKVAVDTTMRDGHQSLLATRVTRTHDIARVASVSKALPOLLSECGWGTDFVSMRFL 616
591 FEDPMDRLDELREAMNNVNIOMLLRGRNTVGVTPVDSVCRAFKVKEAASSGVDIFRIFDA 650
617 TEDPHERLSLREGAPNLLQMLLGGANGVGNVNDVVKYFVQAAAGGVDLFRVFD 676
651 LNDVSMRPAIDAVLETTNATAVAEMAYSGBLSDPNEKLYTLDYLLKMAEIVKSGAHIL 710
677 LNWVENRVSMDAIAEEN-KLCEATICTYTGOLLNSAREPKYDKYTNLAVLEKAGAHII 735
711 AJKDMAGLLRPAATVKLVTALRRPDLFVHVHTHTDAGGQLATVFAAQAAGADAVDGASA 770
736 AVKDMAGLLKPAAKVLPKALREATGLFHFHTHTDTSIGISAATVLAADVADDAAMD 795
771 PLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRV 830
796 AFSGNTSQPCLSGIVEALSSESDTGLDTEWIRRIISFYWEAVRQYAAAFESDLKGPASEV 855
831 YRHEIPGGQSLNRAQATLGLADRELIEDNYAANEMLRGPTKVTPESSKVGDLALHL 890
856 YLHEMPGGQFTNLKQASLGLSEHVEVAQYADANRMFGDIVKVTPESSKVGDMALMM 915
891 VGAGVDPADPAADPOKYDIPDSVIAFLRGLGNPFGWPEPLRTRALRSEKAPLITEV 950
916 VSQDLTVADVENPREVSPDSVSMKGLGQSPGWPEALQKAL----KGEKPYTVR 971
951 PEE--EQAHLDADDSDKERRNSLNR-----LLFPKTEEFLEHRRRFGNTSALDDR 998
972 PGSLEDAALDA-ERKVIETKLERKVDDEPFASYLMFKVTFDFALTAETVGPVSLPTH 1030
999 EFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMNRNVANVNGQIRPMVRDRS-V 1057
1031 AVFYGMEDGEELFADIERGKTLVIVN-QASSGIDDKGMVTVFEINGQPRRIKVPDRAHG 1089
1058 EVTATAEAKDSSNKGHVAAFPAGVVT-VTVAGDEVKAGDAVAIIEAMKWEATITASVD 1116
1090 ASGSAVRKAEPGNASHIGAPMGVLSRVFINQGEVKGAGDVLLSIEAMKWEATITASVD 1149
1117 GKIDRVVPAATKVEGGDLIV 1137

Db 1150 GKIAEVLVKGQIDAKDLI 1170

RESULT 13

AE3285

pyruvate carboxylase (EC 6.4.1.1) [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C;Accession: AE3285

R;DalVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova, R.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesse, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: AE3285

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1158 <KUR>

A;Cross-references: GB:AE008917; PIDN:AAL51448.1; PID:gi7982157; GSPDB:GN00190

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEI0266

A;Map position: 1

A;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

C;Keywords: ligase

Query Match 42.6%; Score 2464.5; DB 2; Length 1158;

Best Local Similarity 46.6%; Pred. No. 1-2e-127;

Matches 538; Conservative 171; Mismatches 408; Indels 37; Gaps 15;

Qy 14 KILVANGETAVRAFAALTEGTAATVAIYPREDGSHFHSFASAVRIG-----TEGSP 67

Db 10 KILVANRESEIAIRVFAANLGLTKTIVAEEDKLSLHFRKADSEYQVGRGPHLDRLGP 69

Qy 68 VKAVLDIDELIIGAAGKVKADAIYPGFLSENAQLARECAENGITIGTPTPEVLDLTDGK 127

Db 70 IESLSIDEIIRVAKUSGADAIHPGGLLSSEFEFEACENGIIVIGKPKETMRRLGNK 129

Qy 128 SRVATRAAKAGLPVLAESTP-SKNIDEIVKSAEGQYPIPVKAVAGGGGRMRFVASPDE 186

Db 130 VAARNLAIEIGVPVVPATDLPDDMDKVAQIGVPLMKASWGGGRGMEATRAEAD 189

Qy 187 LRKLTASAEAEAFGGGAVYVERAVINPHIEVQLDHTGVVHLVERDCSLORRHO 246

Db 190 LAREVMEAKREAAFGKDEYLEKLVERRARHVEVQLGDTYGNVHLFERDCSIRRNQ 249

Qy 247 KVVETIAPAOHLPELRDRIACADAVKFCRSICYGAGTVEFLVD-EXGNHVFIEIMPRIQV 305

Db 250 KVERAPAPYLNDQAQRRELADYGLKIAHADTYIGAGTVEFLMDADTKGFYEVNPRIOV 309

Qy 306 EHTVTEVTEVDLVKAQRLAAGATL--KELGL-TQDKIKTHGAALOCRTTDEPNNGFR 362

Db 310 EHTVTEVTGDIYKAIHILEGPAIGTPEGVPRQEDIRLNGHALQCRITTEDPQNFI 369

Qy 363 PDTGTITAYRSPGGAGVRLD-GAQLGGEITAHFDSMLVMTCRGSDFFETAVARAQALA 421

Db 370 PDYGRIOAYRSAAGFIRLDGGTAYSGAFITRYVDPLLVKTATAGTPELAIHMDRALR 429

Qy 422 EFTVSGVATNIGFIRALLREEDFTSKRIATGFIADHPHLQAPPADDEQRIIDYADVT 481

Db 430 EFRIRGVATNLITPEAIINHFKLSNDYITFRIDTTPPEFQMKRQDRATKLTLYADVT 489

Qy 482 VN-----KPHGVKPKDAAPDKLPNTKDLPLPRGSDRLKQLGPAFAFARDLREQDALV 536

Db 490 VNGHPETKRAKPAADAAP--RVVPFGDKLVADGTQKLLDQLDQPKFAEWMEKEALI 547

Qy 537 TDTTFRDAHQSLIATRVSPALKFAEAAVAKLTPELLSVANGGATVDVAMRFLFEDPWD 596

Db 548 TDTTMRDGHQSLIATRVYTDIARIANAQAALNLFSLCEWGGATFDVSMRFLTEDPWE 607

Qy 597 RLDELREAMPNVNQLMRGNTVGYTPYDPSVCRFAFKVEAASGVDFIFRFDALNDVSQ 656

Db 608 RLALVREGAPNLLQMLRGANGVGYKSPDNNVVKYFVREARAGIDLFRVFDLSLNVEN 667

Qy 657 MRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAREIVKSGAHLAIKDMA 716

Db 668 MRVMDAVLEEN-KICEAAICYTGDILNPDRAKYDLNYYVNLAKYEVKAGAHIIIVKDMA 726

Qy 717 GLLRPAAVTKLVTALRRREFDLPVHVHTHTAGQCLATYFAAAQAGADAVDGASAPLSGTT 776

Db 727 GLLKPAARVLFKALREETDLPPIHFTHTDTSIGISAATVLAADAGVDVDAAMDALSGNT 786

Qy 777 SQPSLSAIVAAFAHTRDRTGLSLEAVSDLEFPYHEAVRGLYLPESGTPGPTGRVVRHEIP 836

Db 787 SQPCGLSIVBALHGSERDGLDPLIRISFYFEAVRHQVAAPFESDLKGPASEVYLHEMP 846

Qy 837 GGQLSNLRQAATLGLADRLDELIEDNYAANEMGLRPTKVTSPSSKVVGD/LALHLVGAGVD 896

Db 847 GGQTNLKEQARSISLGLTRWHEVAQYADVNRMGDIVKVTSPSSKVVGD/MALMMVADLT 906

Qy 897 PADPAADPKQYDIPDSVIAFLRGELGNPPGCGWPEPIRTRALESRSEKAPLITEVPEE-E 954

Db 907 VADVENPAKDIAFPDPSVVMRGDLGQPPSGNPEALQKKYL----KDEKFTTVRPGSLLP 962

Qy 955 QAHLDA-----DQSKERRNS-----LNRLLFPKPTFEHRRRFGNTSALDDREFFYGLV 1005

Db 963 AADLDAERKSFEDSVGRKLSQDFASALMYPKVFTDYATAHETVGTPTVLPTPVYFYLK 1022

Qy 1006 EGRETILRLPDVTRPLLRDLDAISEPDDKGNRVNANVNGQIRPMRVDRDRSVSVATAE 1065

Db 1023 PEEEVFVLDLGRGKTLVIN-QAMSETDEKGMVTVPFELNGQPRRIKVPNRAKASGGVRR 1081

Qy 1066 KADSSNGHVAAPAGVY-TVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVV 1124

Db 1082 KVEAGNDKQVAGMPGVISVAVVAGQVTKGVDLLSIEAMKMETAIHAERDGTIAEVLV 1141

Qy 1125 PAATKVEGGLIVV 1138

Db 1142 RFGQIDAKOLLIV 1155

RESULT 14

T20346

pyruvate carboxylase (EC 6.4.1.1) D2032.2 [similarity] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000

C;Accession: T20346

R;Kershaw, J.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19260

A;Accession: T20346

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1175 <WTL>

A;Cross-references: EMBL:Z81052; PIDN:CAB02872.1; GSPDB:GN00023; CESP:D2023.2

A;Experimental source: clone D2023

C;Genetics:

A;Gene: CESP:D2023.2

A;Map position: 5

A;Introns: 34/3; 103/3; 246/1; 451/3; 567/2; 821/1; 1045/3; 1126/3

C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

C;Keywords: ligase

F;1140/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 42.6%; Score 2464; DB 2; Length 1175;

Best Local Similarity 45.2%; Pred. No. 1.4e-127;

Matches 523; Conservative 200; Mismatches 389; Indels 46; Gaps 14;

Qy 12 FKILVANREGETAVRAFAALTEGTAATVAIYPREDGSHFHSFASAEAVRIGTEGSPVKAY 71

Db 32 FKNVWVANGELAIKRVFALTELNKTYSVAIYAEQDKSNVHLKADAEYLVGKGLPPVAAY 91

Qy 72 LDIDSIIGAAKKVADAIYPGFLSENAQLARECAENGITIGTPTPEVLDLTDGKSAV 131

Db 92 LTIQIIEITALKHNDIAIHGPGFLSERSDFAAACQNAQIGVIFGSPDVMAMGDKVAAR 151

Qy 132 TAAXKAGLPVLABST-PSKNIDEIVKSAEGQYPIPVKAVAGCGGGRMRFPVSPDDELRL 190
 Db 152 QAAIEAGVVPVPGTPTADEAVEFAKQYGTPIILKAAYGGGGRGRRVVDKLVEVEEA 211
 Qy 191 ATASAEAAAGGAGVAVVERAVINPOHLEVOILGDHTEGVHLYERDCSLQRHOKVVE 250
 Db 212 FRASYEAQAAGGDSGLFVEKFERPHIEVLQGDHGNVHLHYERDCSVQRHOKVVE 271
 Qy 251 IAPAQHLDELDRICADAVKFCRSIGYAGTAVFLVDEKGNHVFIEKNPRIQVEHTVT 310
 Db 272 IAPAPALPEGVREKILADALRLARHVGYAGTAVFLVDQKNYVYFIEVNLARLQVEHTVT 331
 Qy 311 EEVTEVDLVKAQVRLAAGATLKELGTQDKIKTHGAALQCRITTEDPNNRPPDGTITA 370
 Db 332 BEITGVDLVQAQRIAGSLDOLKLSQTIQTTSALQCRITTEDPNNRPPDGTITA 391
 Qy 371 YRSPGAGVRLDGAALQGE-ITAHFDSMLVMTCRGSDFTAVARAQALAEFTVSGVA 429
 Db 392 FRSGEGMIRLDSASAFAGSVISPHYDSLWVVIASARNHPNAAAKMIRALKPKRIRGVK 451
 Qy 430 TNGELRALREDEFTSKRTIAGFIADHPLHLCAPADDEQGRILDLADLVYVTKP- 485
 Db 452 TNGELRALREDEFTSKRTIAGFIADHPLHLCAPADDEQGRILDLADLVYVTKP- 511
 Qy 486 -HGVREKDVAAIDKLPNKILPLPRGSRDRILKQGPAAFAFARDLREQDALAVTDTTFRDA 544
 Db 512 ATDLKEAVVSPPIPIYIP--AGAKPTGLRDVLVQRPTEFAKEVRSRPGCMITDTTFRDA 569
 Qy 545 HQSLLATRVRSFALKPAEAAVAKLPPELLSVAGWAGATVDVAMRFLFEDPDRLDELREA 604
 Db 570 HQSLLATRVRTVDMMAISFVQAQSNGLFSLNMGWATPDVSMRFLHCPMERLQTLRL 629
 Qy 605 MPNVNTQMLLRGNTVGYTPYDPSVCRFAVKEAASGVDFIPFIDALNDVDSOMRPAIDAV 664
 Db 630 IPNIPFOCLLRGANMGVSNYPDNVYKFCELAVXNGMDVPRVFDLSVNLVNLVMEAV 689
 Qy 665 LETNTAVAVAMAYSGDSDPNKLYTLDYIKMAEIEVKSAHILAKWAGLLRPAAV 724
 Db 690 GRAG-GVFAATATYDGVTDKRDYDLKYINLADQLVKAQAHILSKWAGVLPKPA 748
 Qy 725 TKLVTLARREF-DLPVHVHTHTAGGLATYFAAQAGADAVDGAAPLSTGTSQPSLSA 783
 Db 749 KLLIGALRKDFDIPHIVHTHTDSGAGVAAMLEKAGADVDAVDSWGSMTSOPSWGA 808
 Qy 784 IVAAFATRRDGLSLEAVSDLEPYWEAVRGVLYLPFESGTPGPTGR--VYRHEIPGGQLS 841
 Db 809 IVASLQGTQKHTDGLSLDDISKYSAYWESTRLYAPFECATTKWKSQNVADYKHEIPGGQYT 868
 Qy 842 NLRAQATLGLADRFELIEDNYAAVNMELGRPTKVTTPSSKVVDLALHLVGAG-----V 895
 Db 869 NLQFQAFSLGLGQPFDEVKEMREANLVLDIIVKTPSSKIVGDLAQFWQNNLTRETLV 928
 Qy 896 DPADFAADPKQYDIPDSVIAFLRGELGNPPGGNPPPLRTRALEGRS--EGKAPLVEPEE 953
 Db 929 DRAD-----DLSPFKSVVDVMQNVGQPPYGFPEPLRTKVLRGKPKVDGR-----PGE 976
 Qy 954 EQAHLADDSK-----ERNNSNRLFLFKPTEEFLEHRRFNGNTSALDDREFFY 1002
 Db 977 NAKPVDLDVAKVELEKHEKRTISEEDVMYSNFTPTVDFEFTRQOYGPVCLKPLRLFLT 1036
 Qy 1003 GLVEGETILIRPDVPTLLRLDAISEPDDKGMNVANVNGQIRPMRVEDRSVESVTA 1062
 Db 1037 GLIEAEVDVEIESGKT-LAIQLLASGKLNKRGEREVEFPDLNGQKRSIFVVDKESKEIV 1095
 Qy 1063 TAKKADSSNKHVVAAPFAG-VVTVTVVAEGDEYKAGDAVAIIEAMKMEATITASVDGKIDR 1121
 Db 1096 TRPALPGVGRHIGAMPDGVDELKIKEGDKGVTKKQPLFVLSAKMNMWIDSPAGTVKA 1155
 Qy 1122 VVYPAATKVEGGDLIVV 1139
 Db 1156 IHAPQGTCKSAGDLVVEV 1173

RESULT 15
 QYBP
 pyruvate carboxylase (EC 6.4.1.1) 1 [validated] - yeast (Saccharomyces cerevisiae)
 N;Alternate names: protein G3428; protein YGL062w; pyruvic carboxylase
 C;Species: Saccharomyces cerevisiae
 C;Date: 31-Dec-1991 #sequence-revision 19-Jul-1996 #text_change 01-Feb-2002
 C;Accession: S64066; A29233; S05760; A29722
 R;Feuermann, M.; Poirier, S.; Souciet, J.L.
 submitted to the Protein Sequence Database, May 1996
 A;Reference number: S64044
 A;Accession: S64066
 A;Molecule type: DNA
 A;Residues: 1-1178 <FEU>
 A;Cross-references: EMBL:Z72584; NID:gl322565; PIDN:CAA96765.1; PID:gl322566; GSPDB:GNO
 R;Lim, F.; Morris, C.P.; Occhiodoro, F.; Wallace, J.C.
 J. Biol. Chem. 263, 11493-11497, 1988
 A;Title: Sequence and domain structure of yeast pyruvate carboxylase.
 A;Reference number: A92662; MUID:88298805; PMID:3042770
 A;Accession: A29233
 A;Molecule type: DNA
 A;Residues: 1-461, 'G', 463-492, 'D', 494-594, 'A', 596-618, 'Q', 620-663, 'S', 665-771, 'R', 773-8
 A;Cross-references: EMBL:J03889; NID:gl172101; PIDN:AAA34843.1; PID:gl172102
 R;Morris, C.P.; Lim, F.; Wallace, J.C.
 Biochem. Biophys. Res. Commun. 145, 390-396, 1987
 A;Title: Yeast pyruvate carboxylase: Gene isolation.
 A;Reference number: S05760; MUID:87241529; PMID:3036126
 A;Accession: S05760
 A;Molecule type: DNA
 A;Residues: 1003-1178 <MOR>
 A;Cross-references: EMBL:J03889
 A;Accession: A29722
 A;Molecule type: protein
 A;Residues: 1124-1149 <MOR2>
 C;Genetics:
 A;Gene: SGD:PYC1; PVY; MIPS:YGL062w
 A;Cross-references: SGD:S0003030; MIPS:YGL062w
 A;Map position: 7L
 C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
 C;Keywords: biotin binding; gluconeogenesis; homotrimer; ligase; zinc
 F:21-478/Domain: biotin carboxylase homology <BCH>
 F:157-331/Domain: ATP/bicarbonate binding #status predicted <ATB1>
 F:353-468/Domain: ATP/bicarbonate binding #status predicted <ATB2>
 F:569-908/Domain: pyruvate binding #status predicted <PYK>
 F:1096-1169/Domain: lipoyl/biotin-binding homology <LPB>
 F:1135/Binding site: biotin (lys) (covalent) #status experimental
 Query Match 42.4%; Score 2457; DB 1; Length 1178;
 Best Local Similarity 46.7%; Pred. No. 3 3e-127; Mismatches 44; Gaps 16;
 Matches 54; Conservative 173; Indels 401; Indels 44; Gaps 16;
 Qy 14 KILVANRGEIATVAPRAALETGAATVAIYPRDRGSHRSFASAVRIGTEG--SPVKAY 71
 Db 21 KILVANRGEIPIRIFRTAHELSMCTVAIYSHEDRLSTHRKQKADAYVIGEVGYTPVGAY 80
 Qy 72 LDIDEIIGAARKVKADAIYVGYGFLSENACLAECENGTITFTGPTPEVLDTLGDKSRV 131
 Db 81 LAIDEIISIAQKQVDFIHPGYGFLSENSEFADKVKVAGITWIGPFAEVIDSDGVKVSAR 140
 Qy 132 TAAKKAGLPVLABST-PSKNIDEIVKSAEGQYPIPVKAVAGCGGGRMRFPVSPDDELRL 190
 Db 141 NLAAKANVTPVPTGPIETVEEALDFVNEYGYPIVIAAFGGGGRGMRVREGDDVADA 200
 Qy 191 ATASAEAAAGGAGVAVVERAVINPQHLEVOILGDHTEGVHLYERDCSLQRHOKVVE 250
 Db 201 FORATSEARTAFNGTCFVERFLDKPKHIEVQLADNHNHNVHLFDRDCSVQRHOKVVE 260
 Qy 251 IAPAQHLDELDRICADAVKFCRSIGYAGTAVFLVDEKGNHVFIEKNPRIQVEHTVT 310
 Db 261 VAPAKTLPREVRDAILTDAVKLAKECGYRNAGTAFAELVDNQNHRHYFIEINPRIQVEHTIT 320
 Qy 311 EEVTEVDLVKAQVRLAAGATLKELGTQDKIKTHGAALQCRITTEDPNNRPPDGTITA 370

```

321 EEITGIDIVAAQIQIAGASLPQGLFQDKITTRGFQAIQCRITTEDPAKNFQDGTGRIEV 380
371 YRSPGGAGVLDGA-AQLGGEITAHFDSMLVKMTCRGSDPFETAVARQAALAEFTVSGVA 429
381 YRAGNGVLDGNGAYAGTIISPHYDSMLVKCSGSGTYEIVRRKXIRPALLIEFIRGVK 440
430 TNGFRLALREEDFTSKRIATFIADHPLHQAAPPADDEQGRILDYLDADVTWN---KP 485
441 TNIPFLLTLTNPFVTEGTWTFIDDTPOLFQWSSQNRQAQKLLHLYLADVAVNGSSIKG 500
486 HGVPRDVAAPIDKLNKDL-----PLPRGRDRDLKOLGPAAFAARDLREODALA 535
501 QIGLPLKXNP--SVPHLHDAQGNVINTKSPPSGWRQVLLKGPAPFAQVRQFNGTL 558
536 VTDITFRDAHQSLLATVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPW 595
559 LMDTTWRDAHQSLLATVRSFTHLTIAPTTHALAGRFALCWCWGATFDVAMRFLHEDPW 618
596 DRLEDEAREAMPNINIOMLGRNTVGYTPYPSVCHAFVKEAASSGVDFRIFDALNDVS 655
619 ERUKRSLVFNIPFQMLLRGANGVAYSSLPDPAIDHVFVKQAKNGVDIFRVFDALNDLE 678
656 QMRPAIDAVLETTNTVAEAVAMAYSGDLSDFNEKLYTLDYLLKMAEIVKSGAHILAIDKM 715
679 QLVGVDAVKKAG-GVVEATVCFSGDMLQPGKK-YNLDYVLEIAEKIVQMGTHILGIKDM 736
716 AGLLRPAATVKTLYTALREF-DLPVHVHTDHTAGGQATYFAAAQACADAVDCASAPLSG 774
737 AGTKPAAAKLLIGSLRAKYPDPIHVHTHDSAGTAVASMTACALAGADVVDVAINSMG 796
775 TTSOPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFBESGTPPGTGRVYRHE 834
797 LTSOPSNALLASL-EGNIDTGINVEHVRELDAYWAEMRLLYSCFEADLKGPDPPEVYQHE 855
835 IPGQLSNLRAQATAGLADRPFLIEDNYAAVNEMLGRPTKVTTPSSKYVGDALHLVVGAG 894
856 IPGQQLTNLLFQAQQLGLGEQWAEKRAYREANYLLGDIIVKVTPTSKYVGDLAQFMYSNK 915
895 VDPADPAADPKYDIPDSVIAFLRGLGNPPGQWPEPLRTRALEGRSEKGAPLTEVPEEE 954
916 LTSDDVRLANSLDFPDSVMDFFEGGLIGQYGGFPFPPFRSDVLENK---RRKLTCPGLE 972
955 QAHLDDDDSKERNNSLNL-----LFPKTEEFLEHRRRFGNTSALDDREFFY 1002
973 LEPPDLE--KIREDLQNRFGVDDECDVASYNMYPRVYEDFQKMETYGDLSVLPTRSFLS 1030
1003 GLVEGRETLRLPDVRLPRLVRLDAISEPDK-GMRNVVANVANGJIRPMRVRDRSVESVT 1061
1031 PLETDEIEVIEGQKT-LIIKQAVGDLNKKTGEREVYFDLNGEMRKIRVADRQKXVET 1089
1062 ATAEKADSSNKGHVAAPFAGV-VTVTVAGDEVKAGDAVAIIEMXKHEATITIASVDGKID 1120
1090 VTKSKADMDHPLHIGAPMAGVIVEVKVHKSGLIKKQPVAVLSAMKMEMIISSPSDQGVK 1149
1121 RVVVPATKYEGGDLIVVV 1139
1150 EVFVSDGENVDSDLLVLL 1169

```

Search completed: March 24, 2004, 22:44:53
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:36:41 ; Search time 17 seconds
(without alignments)
3491.763 Million cell updates/sec

Title: US-10-045-072-2

Perfect score: 5788
Sequence: 1 MSTHTSSTLPAPKKILVANR.....RVVVPAAKVEGGDLIVVWS 1140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2541.5	43.9	1178	1 PYC_MOUSE	Q05920 mus musculus
2	2525.5	43.6	1178	1 PYC_RAT	P52873 rattus norv
3	2520.5	43.5	1178	1 PYC_HUMAN	P11498 homo sapien
4	2457	42.4	1178	1 PYC1_YEAST	P11154 saccharomyc
5	2436	42.1	1180	1 PYC2_YEAST	P32327 saccharomyc
6	2391	41.3	1189	1 PYC_PICPA	P78992 picchia past
7	1079	18.6	501	1 PYC1_METJA	Q58526 methanococc
8	1036.5	17.9	447	1 ACCC_ANASP	Q06862 anabaena sp
9	1019	17.6	506	1 PYCA_ARCFU	Q30019 archaeoglob
10	997	17.2	491	1 PYCA_METTH	O27939 methanobact
11	946	16.3	448	1 ACCC_HABIN	P43873 haemophilus
12	938	16.2	449	1 ACCC_ECOLI	P24182 escherichia
13	936	16.2	449	1 ACCC_ECO57	Q8x9b6 escherichia
14	924	16.0	449	1 ACCC_PSEAE	P31798 pseudomonas
15	911	15.7	725	1 MCCA_HUMAN	Q86rq3 homo sapien
16	909	15.7	717	1 MCCA_MOUSE	Q99mr8 mus musculu
17	883.5	15.3	654	1 BCCA_MYCTU	P46401 mycobacteri
18	879	15.2	448	1 ACCC_BACSU	P49787 bacillus su
19	875.5	15.1	598	1 BCCA_MYCLE	P46392 mycobacteri
20	871.5	15.1	567	1 PYCB_METJA	Q58628 methanococc
21	866	15.0	703	1 PCCA_HUMAN	Q05165 homo sapien
22	859.5	14.8	731	1 MCCA_SOYBN	P42177 glycine max
23	859	14.8	734	1 MCCA_ARATH	Q42523 arabidopsis
24	843.5	14.6	704	1 PCCA_RAT	P14882 rattus norv
25	835.5	14.4	1835	1 DURL_YEAST	P32528 saccharomyc
26	825	14.3	568	1 PYCB_METTH	O27179 methanobact
27	804	13.9	590	1 DCOA_SALTY	Q03030 salmonella
28	799.5	13.8	595	1 DCOA_KLEBN	P13187 klebsiella
29	612.5	10.6	2345	1 COAL_RAT	P11497 rattus norv
30	609.5	10.5	2324	1 COAC_CHICK	P11029 gallus gall
31	609.5	10.5	2346	1 COAL_HUMAN	Q13085 homo sapien
32	608.5	10.5	2346	1 COAL_BOVIN	Q9tts8 bos taurus
33	601	10.4	2346	1 COAL_SHEEP	Q28559 ovis aries

ALIGNMENTS

RESULT 1

ID	PYC_MOUSE	STANDARD	PRT	1178 AA
AC	Q05920;1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic carboxylase) (PCB).			
GN	PC OR PCX			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	SEQUENCE FROM N.A.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Adipocyte;			
RX	MEDLINE=93189578; PubMed=8446588;			
RA	Zhang J., Xia W.L., Brew K., Ahmad F.;			
RT	"Adipose pyruvate carboxylase: amino acid sequence and domain structure deduced from cDNA sequencing."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:1766-1770(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FVB/N; TISSUE=Liver;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,			
RA	Butterfield A.S., Krzywinski M.I., Skalska U., Smallos D.E.,			
RA	Schmerch J., Schain J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-!- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction, involving the ATP-dependent carboxylation of the covalently attached biotin in the first step and the transfer of the carboxyl group to pyruvate in the second. Catalyzes in a tissue specific manner, the initial reactions of glucose (liver, kidney) and lipid (adipose tissue, liver, brain) synthesis from pyruvate.			
CC	-!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3) (-) = ADP + phosphate + oxaloacetate			
CC	-!- COFACTOR: Biotin and manganese			
CC	-!- PATHWAY: Gluconeogenesis and lipogenesis.			

C00763 homo sapien
P32874 saccharomyc
Q00955 saccharomyc
P78820 schizosacch
O28994 archaeoglob
Q27077 methanobact
Q8x338 escherichia
P00968 escherichia
Q8dgm2 vibrio vuln
Q9w227 thermotoga
P14846 salmonella
Q87sf3 vibrio para

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RC MEDLINE=96096548; PubMed=8522203;
 RA Lehn D.A., Moran S.M., Macdonald M.J.,
 RL "The sequence of the rat pyruvate carboxylase-encoding cDNA.";
 RN Gene 1651331-332(1995).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Wistar; TISSUE=Liver;
 RC MEDLINE=96257760; PubMed=8687410;
 RA Jitrapakdee S., Booker G.W., Cassidy A.I., Wallace J.C.;
 RL "Cloning, sequencing and expression of rat liver pyruvate
 carboxylase.";
 RN Biochem. J. 316:631-637(1996).
 CC -!- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
 involving the ATP-dependent carboxylation of the covalently
 attached biotin in the first step and the transfer of the carboxyl
 group to pyruvate in the second. Catalyzes in a tissue specific
 manner, the initial reactions of glucose (liver, kidney) and lipid
 (adipose tissue, liver, brain) synthesis from pyruvate.
 CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3) (-) = ADP + phosphate +
 oxaloacetate.
 CC -!- COFACTOR: Biotin and manganese (By similarity).
 CC -!- PATHWAY: Gluconeogenesis and lipogenesis.
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sb-sib.ch).
 CC
 DR EMBL; U32314; AAA96256.1; -;
 DR EMBL; U36585; AAC52668.1; -;
 DR PIR; S68252; JC4391.
 DR HSRF; P24182; IENC.
 DR InterPro; IPR001882; Biotin_BS.
 DR InterPro; IPR005482; Biotin_carb_C.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR005479; Cpsase_L_D2.
 DR InterPro; IPR005481; Cpsase_L_N.
 DR InterPro; IPR000891; HMGL-Like.
 DR InterPro; IPR003379; PVC_OADA.
 DR InterPro; IPR005930; Pyruv_carbox.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF00364; Biotin_lipoyl; 1.
 DR Pfam; PF00289; Cpsase_L; 1.
 DR Pfam; PF02786; Cpsase_L_D2; 1.
 DR Pfam; PF00682; HMGL-Like; 1.
 DR Pfam; PF02436; PVC_OADA; 1.
 DR TIGRfam; TIGR01235; Pyruv_carbox; 1.
 DR PROSITE; PS00188; Biotin; 1.
 KW Ligase; Multifunctional enzyme; Lipid synthesis; Transit peptide;
 KW ATP-binding; Mitochondrion; LIPOAMIDE TRANSFERASES;
 FT TRANSIT 1 20
 FT CHAIN 21 1178
 FT DOMAIN 21 549
 FT DOMAIN 550 1000
 FT DOMAIN 1096 1178
 FT NP_BIND 198 203
 FT ACT_SITE 328 328
 FT BINDING 1144 1144
 FT BINDING 222 222
 FT CONFLICT 222 222

FT CONFLICT 866 866 D -> I (IN REF. 2).
 FT CONFLICT 977 977 G -> R (IN REF. 2).
 SQ SEQUENCE 1178 AA; 129689 MW; 8B5FAL9BC132A8DD CRC64;
 Query Match 43.6%; Score 2525.5; DB 1; Length 1178;
 Best Local Similarity 46.4%; Pred. No. 1.9e-127;
 Matches 541; Conservative 185; Mismatches 410; Indels 31; Gaps 13;
 QY 1 MSTHTSSTLPA-----PKILVANRGEIAVRAFAALETGAATVAIIPREDRG 48
 DB 14 LGVRSSTAPVSPNVRLEVKPIKKVMVANRGEIAIRVFRACTELGIKIRTVAVVSEDTG 73
 QY 49 SFHRSFASAVRIGTEGSPVKAYIDIDEITCAAKVKVADAIYPCYGLSENAQLARCAE 108
 DB 74 QMHRQKADAEAYLIGRLAPVQAYLHPDIIVKAKENGVDVHPGCGYGLSRAADPAQACD 133
 QY 109 NGITFIGPTPEVLDITGDKSRAVTAAGKAGLPVL-AESTPSKNIDETVKAEGQTVIFV 167
 DB 134 AGVRFIPGSPPEVRMGDKVEARAIAAAGVVPVFGTNSPINSIHEAHSNTVGFPIIF 193
 QY 168 KAVAGGGGGRGMRFVASPELAKLATEASREAAAFGAGVYVERAVINPQIHIEVQIIGDH 227
 DB 194 KAAVGGGGGRGMRVHVSVEELEENYTRAYPEALAAFGNGALFVEKFIKPRHIEVQIIGDQ 253
 QY 228 TGEVHLVERDCSLQRHOKVVEIAPAGHLDPELRDRIADAVKFCRSIGQAGTVEFL 287
 DB 254 YGNILHLVERDCSLQRHOKVVEIAPATHLPQILSRILTSVSKLAKGVYENAGTVEFL 313
 QY 288 VDEKGNHYFIEMNPRIQVEHTVTVEEDLVKQMLAAGATLKEGLTQDKIKTHGAA 347
 DB 314 VDKGHKHYFIENSRLOVEHTVTEITDVLVHAQIHVSEGRSLPDLGLRGENIRINGCA 373
 QY 348 LQCRITTTDDPNNGPRPDGTITAYRSFGAGVRLDGAQL-GGEITAHFDSMLVMTCRG 406
 DB 374 IQCRVTTDDPARSPQDTRIEVFRSGEGMIRLDNAAFGAVISPHYDILLVKVIAHG 433
 QY 407 SDFETAVARAQALAEFTVSGVATNIGFRLALLREEDFTSKRIATGFTADPHLLQAPPA 466
 DB 434 KDHPATAAKMSRALAEFRVGRKINI PFLQNVLNNOQLAGIVDTQFIDENPELQLRPA 493
 QY 467 DEQGRILDYLDVTVNKGVRPKDVA-APIDKL-PNIKOLPLPGSGEDRLKQLGPAAF 524
 DB 494 QNRAQKLLHYLGHVMVNGFTTPIPVKVSPPVDPIVPPVPGPPAGPRDILLRGGPGF 553
 QY 525 AEDLREQDALAVDTTFRDAHQSLIATVRSPALKAEEAVAKLTPELLSVEANGATYD 584
 DB 554 ABAVRNHQGLLLMDTTFRDAHQSLIATVRTHDLKKIAPYVAHNFNFLFSIENWGGATFD 613
 QY 585 VAMRFLFEDPDWRDLDELREAMPNNIOMLLRGNTVGTPTPDSVCRFAVKEAASGVDI 644
 DB 614 VAMRFLYECPMRRLQELRELIPNIPFQMLLRGANAAGVTNPDNVVFKFCEVAKENGMDV 673
 QY 645 FRIFDALNDVSOMRPAIDAVLENTNTAVAEVAMAYSGDLSDPNEKLYTLDYVYLMKAAEIVK 704
 DB 674 FRIFSLNLYLPMILGMAAGSAG-GVVEAALSYTGDVADFSRTKYSLEYWGLAEELVR 732
 QY 705 SGAHILAIKDMAGLLRPAAVTKLVLTALRREP-DLPHVHTHTTAGQLATYFAAQAGAD 763
 DB 733 AGTHILCIKDMAGLLKPAACCTMLVSLRDRPDLPLHTHTDTSGSGVAAMLACQAGAD 792
 QY 764 AVDGASAPLSGTTSPSLSAIVAAFAHTRRDRTGLSLVSLDLEPVEYAEVARGVLYLPESGT 823
 DB 793 VDVAVDSVSGMTSPQSMGALVACTGTFPDLTEVLEVRFDYSEYEWARGLYAAFDCTA 852
 QY 824 PGPTGR--VYRHFPGGQSLNRAQATAGLADRELIEDNEDYAAVNMGLGRTKVTPSK 881
 DB 853 TMSKSGNSDVYENEDPGGQYTNLHFOAHNSMGLSKFKEVKYAEVANQMLGDIKIVTPSK 912
 QY 882 VVGDIALHLVAGVDPADPAADPKYDIPDSVIAFLRGLGNPPGCGWPEPLKTRALEG-- 939
 DB 913 IVGDIAQFMVQNGLSRAEAEAEELSPRSVVEFLQYIGIPHGSGFPEPFPSKVLKOLP 972
 QY 940 RSEKGAFLTEVPPEEQAHLDAD-----DSKERRNSLNRLLPKPTPEFLEHRRFRGNTS 993

Db 973 RIEG-GPCASLPPLNKLXLEKDLIDREGEVTPEDVLSAAMYDPVFAQKDFDFTATEGPLD 1031
 Qy 994 ALDRBEPYGLVGEETLRLPDVTRPLVRLDAISEDDKGVNNAVNGQIRPVR 1053
 Db 1032 SLNTRFLQGGKIAEFEFELEERKTLHIALKALVSLNIRAGQRFVFEGLQIRSLVX 1090
 Qy 1054 DRSVESVTATAEKADSSNGHVAAPFAG-VVTVTVAEQDEYKAGDAVAIIEAMKMEATIT 1112
 Db 1091 DTQAMKEMHFFPKALKDVKGQIGAPMPGKVIDVKAAGAKVVKGQPLCVLSAMKQETVVT 1150
 Qy 1113 ASVDGKIDRVVPAATKVEGGLIVVV 1139
 Db 1151 SPMEGTIRKRVHTKDMTLEGDLLEI 1177

RESULT 3
 PYC_HUMAN STANDARD; PRT; 1178 AA.
 AC P11498; Q16705;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
 DE carboxylase) (PCB).
 GN PC.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney, and Liver;
 RX MEDLINE=95002202; PubMed=7918683;
 RA Wexler I.D., Du Y., Lisgaris M.V., Mandal S.K., Freytag S.O.,
 RA Yang B.-S., Liu T.-C., Kwon M., Patel M.S., Kerr D.S.;
 RT "Primary amino acid sequence and structure of human pyruvate
 RT carboxylase.";
 RN Biochim. Biophys. Acta 1227:46-52 (1994).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=94324922; PubMed=8048912;
 RA Mackay N., Rigat B., Douglas C., Chen H.S., Robinson B.H.;
 RT "cDNA cloning of human kidney pyruvate carboxylase.";
 RL Biochem. Biophys. Res. Commun. 202:1009-1014 (1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney, and Liver;
 RA Walker M.E., Jitrapakdee S., Val D.L., Wallace J.C.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=22368257; PubMed=12477932;
 RA Strausberg R.L., Faigold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner R., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Muray K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [5]
 RP SEQUENCE OF 1083-1178 FROM N.A.
 RX MEDLINE=87212051; PubMed=3555348;
 RA Lambonwah A.-M., Quan F., Gravel R.A.;
 RT "Sequence homology around the biotin-binding site of human propionyl-
 RT CoA carboxylase and pyruvate carboxylase.";
 RL Arch. Biochem. Biophys. 254:631-636 (1987).
 RN [6]
 RP SEQUENCE OF 1135-1178 FROM N.A.
 RX MEDLINE=8503380; PubMed=6548474;
 RA Freytag S.O., Collier K.J.;
 RT "Molecular cloning of a cDNA for human pyruvate carboxylase.
 RT Structural relationship to other biotin-containing carboxylases and
 RT regulation of mRNA content in differentiating preadipocytes.";
 RL J. Biol. Chem. 259:12831-12837 (1984).
 RN [7]
 RP VARIANTS PC DEFICIENCY THR-610 AND ILE-743.
 RX MEDLINE=98254451; PubMed=9585612;
 RA Carbone M.A., Mackay N., Ling M., Cole D.E.C., Douglas C., Rigat B.,
 RA Feigenbaum A., Clarke J.T.R., Haworth J.C., Greenberg C.R.,
 RA Seagrang L., Robinson B.H.;
 RT "Amerindian pyruvate carboxylase deficiency is associated with two
 RT distinct missense mutations.";
 RL Am. J. Hum. Genet. 62:1312-1319 (1998).
 RN [8]
 RP VARIANTS PC DEFICIENCY ALA-145 AND CYS-451.
 RX MEDLINE=98244401; PubMed=9585002;
 RA Wexler I.D., Kerr D.S., Du Y., Kaung M.M., Stephenson W., Lusk M.M.,
 RA Wappner R.S., Higgins J.J.;
 RT "Molecular characterization of pyruvate carboxylase deficiency in two
 RT consanguineous families.";
 RL Pediatr. Res. 43:579-584 (1998).
 CC -1- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
 CC involving the ATP-dependent carboxylation of the covalently
 CC attached biotin in the first step and the transfer of the
 CC carboxyl group to pyruvate in the second. Catalyzes in a tissue
 CC specific manner, the initial reactions of glucose (liver, kidney)
 CC and lipid (adipose tissue, liver, brain) synthesis from pyruvate.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: Biotin and manganese.
 CC -1- PATHWAY: Gluconeogenesis and lipogenesis.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- DISEASE: Defects in PC are the cause of pyruvate carboxylase
 CC deficiency (PC deficiency) [MIM:266150]. PC deficiency leads to
 CC lactic acidosis, mental retardation and death. It occurs in three
 CC forms: mild or type A, severe neonatal or type B, and a very mild
 CC lactacidemia.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U04641; AAA99537.1; -
 DR EMBL; S72370; AAB31500.1; -
 DR EMBL; U30891; AAA82937.1; -
 DR EMBL; BC011617; AAH11617.1; -
 DR EMBL; M26122; AAA36423.1; -
 DR EMBL; X02282; AAA60033.1; -
 DR PIR; G01933; JC2460.
 DR HSSP; P24182; 1BNC.
 DR Genew; HGNC:8636; PC.
 DR GK; P11498; -
 DR MIM; 266150; -

DR GO: 0005524; F:ATP binding; TAS.
 DR GO: 0009374; F:biotin binding; TAS.
 DR GO: 0004736; P:pyruvate carboxylase activity; TAS.
 DR InterPro: IPR001882; Biotin_BS.
 DR InterPro: IPR005482; Biotin_carb_C.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR005479; Cphase_L_D2.
 DR InterPro: IPR005481; Cphase_L_N.
 DR InterPro: IPR000891; HMGL-Like.
 DR InterPro: IPR003379; PYC_OADA.
 DR InterPro: IPR005930; Pyruv_carbox.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF00364; Biotin_lipoyl; 1.
 DR Pfam: PF00289; Cphase_L_D2; 1.
 DR Pfam: PF02786; Cphase_L_D2; 1.
 DR Pfam: PF00682; HMGL-Like; 1.
 DR Pfam: PF02436; PYC_OADA; 1.
 DR TIGRfam: TIGR01235; Pyruv_carbox; 1.
 DR PROSITE; PS00188; Biotin; 1.
 KW Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconogenesis;
 KW ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide;
 KW Disease mutation.
 FT TRANSIT 1 20 MITOCHONDRION (POTENTIAL).
 FT CHAIN 21 1178 PYRUVATE CARBOXYLASE
 FT DOMAIN 21 549 BIOTIN CARBOXYLASE (BY SIMILARITY).
 FT DOMAIN 550 1000 CARBOXYLTRANSFERASE (BY SIMILARITY).
 FT DOMAIN 1096 1178 BIOTIN CARBOXYL CARRIER PROTEIN
 FT NP BIND 198 203 ATP (BY SIMILARITY).
 FT ACT SITE 328 328 BY SIMILARITY.
 FT BINDING 1144 1144 BIOTIN (BY SIMILARITY).
 FT VARIANT 145 145 V -> A (in PC deficiency; mild).
 FT VARIANT 451 451 R -> C (in PC deficiency; mild).
 FT VARIANT 610 610 A -> T (in PC deficiency; mild).
 FT VARIANT 743 743 M -> I (in PC deficiency; mild).
 FT CONFLICT 225 226 LA -> WP (IN REF. 2).
 FT CONFLICT 352 352 A -> S (IN REF. 3).
 FT CONFLICT 385 386 RS -> PT (IN REF. 2).
 FT CONFLICT 486 487 EL -> DV (IN REF. 2).
 FT CONFLICT 638 638 P -> R (IN REF. 2).
 FT CONFLICT 729 729 E -> A (IN REF. 2).
 FT CONFLICT 774 775 DT -> AP (IN REF. 2).
 FT SEQUENCE 1178 AA; 129633 MW; 381F527553A20095 CRC64;
 Query Match 43.5%; Score 2520.5; DB 1; Length 1178;
 Best Local Similarity 46.3%; Pred. No. 3.4e-127;
 Matches 537; Conservative 183; Mismatches 412; Indels 29; Gaps 12;
 QY 6 SSTLPA-----FKKILVANRGETAVRAFAALETGAATVAIYPRDRGSGFHS 53
 DB 19 TSTAPASPNVRLEYKPIKVVANRGETAIRVFRACTELGIKRTVAIYSEQDTGMHQ 78
 QY 54 FASEAVRIGTEGSPVKAYLDIDEIIGAAKKVADAIYPGYFLSENAQALARECAENGITF 113
 DB 79 KADEAYLIGRLAPVQAYLHPDIKIVAKENNVDAVHPGYFLSERADFAQACQAGVRF 138
 QY 114 IGTPPEVLDTGDKSRATVAKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAG 172
 DB 139 IGPSPVVRKMGDKVEARAIAAGVPVPGTAPITSLHEAFNTGTFPIIFKAAAG 198
 QY 173 GGGGRMFVAPSDRLRLKATEASREAAFGDGAIVYVERAVINPQIEVQILGDHTGEVV 232
 DB 199 GGGGRMVAHVSYELEENYTRAVSEALAFNGALFVEKFEKPRHIEVQILGDQYGNIL 258
 QY 233 HLYEDCSLQRHOKVVEIAPACHLDPELDRICADAVKFCRSIGYCGAGTVEFLVDEKG 292
 DB 259 HLYEDCSIQRRHOKVVEIAPACHLDPELDRICADAVKFCRSIGYCGAGTVEFLVDRHG 318
 QY 293 NHVFIENNPRIQVHTVTEETVDLVKQVLAAGATLKLGLTQDKINTHGAALQCRI 352

DB 319 KHVFIENSLQVHTVTEETVDLVHQAQIHVAEGRSLPDLGLRQENIRINGCAIQCVRV 378
 QY 353 TTEDPNNGPDPDTGTTITAYRSPGAGVRLDGAQL-GGEITAFHSDSLVMTWTCRGSDPET 411
 DB 379 TTEDPNSFPQDTGRIEVFESGEGMGLRDNAGAFQCAVISPHYDSLVLVIANGKHPT 438
 QY 412 AVARAQALAEFTVSGVATNIGFLRALLREEDTTSKRIATGFIADPHLLQAPPADDEQG 471
 DB 439 AATKVSALAEFRVGVKTNIAFLQVNNQQLAGVTDTQFIDENPELFPOLPAQNAQ 498
 QY 472 RILDYADVTVNPKHGVPR-KDVAAPIDK-LPNIKDLPLPGSRDLKQLGPAFAFADLR 529
 DB 499 KLLHYLGHVMVNGPTTPIPKASPSPTDPVVPVAVP:GPPAGPFCILLREGSPGFAVR 558
 QY 530 EQDALAVTDTTFRDAHOSILLATRVSRPALKPAAEVAVALTPELLSVEAWGATYDVAMRF 589
 DB 559 NHPGLLLMDTTFRDAHQSLLATRVTHDLKKIAPYVAHNFSLFSENMWGAFEDVAMRF 618
 QY 590 LFEDPWRDLBELRAMPNNIOMLLRGNVTGVTPTPDSVCRAVFKAEASSGVVDIIFRID 649
 DB 619 LYECFMRRLQELRELIENIPFQMLLRGANAVGVTNYPDNVVFKECEVAKENGMDVRFVD 678
 QY 650 ALNDVSQWRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLTDYLLKMAEEIVKSGAHI 709
 DB 679 SLNLPNKLGMERAGSAG-GVVEAALSYTGDAVDSRTKYSLQYMGLELAEELVRAGTHI 737
 QY 710 LAIKDMAGLRPAATKLVLTALREF-DLPVHVHTHTAGGOLATYPAAQAQADAVDGA 768
 DB 738 LCIKDMAGLLKPTACTMLVSLRDRFPDPLPHIHTHTDTSAGVAAMLACQAQAGADVDDVA 797
 QY 769 SAPLSGTTSTPSLSAIAVAFAHTRDRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTGPTG 828
 DB 798 ADSHSGMTSPQSMGALVACTRGTPDTEVPMERVDYSEWEGARGLYAAFDCTATMKS 857
 QY 829 R--YRHEIPGQUSNLRQAATLGLADRELIEDNYAANVEMLRGRTKVTTPSKVVDGL 886
 DB 858 NSDVEYNEIPGGQYTNLHFQAHSMGLSGKFEVKKAYVEANQMLGDLIKVTPSSKIYVDL 917
 QY 887 ALHLVAGVDPADFAADPKYDIPDSVIAFLRGLGNPPGCGWPEPLRTRAL-----EGR 940
 DB 918 AQFVQNGLSRAEAAQAEELSFPSSVFELOGYIGVPHGFPPEPFSSKVLKDLPRVEGR 977
 QY 941 SEGKAPLVEPPEEQAHLDA-DDSKERNSLNRLFPKPTTEFLEHRRRFGNTSALDDRE 999
 DB 978 PGASLPDLQLQALEKELVDRHGEVTPEDVLSAAMYPDVFAHKDFTATFGPLDSLNTRL 1037
 QY 1000 PFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMENVVANVNGOIRPMRVDRDSVES 1059
 DB 1038 FLOQPKIAEEFEVELEKGT-LHIKALAVSDLNAGORQVFFELNGQLRSILVKDQTAMK 1096
 QY 1060 VTATAEKADSSNKGHVAAFPAG-VVTVTVAEGDEVKAGDAVAITAEAMKMEATITASVDGK 1118
 DB 1097 EMHFPKALKDVKGQIGAPMFGKVIDIKWVAGAKVAKGQPLCVLSAMKMETVVTSPMEGT 1156
 QY 1119 IDRVVVPAAKVEGGDLIVVV 1139
 DB 1157 VRKVHTKDMTLEGGDUILEI 1177
 RESULT 4
 PYC1 YEAST STANDARD; PRT; 1178 AA.
 ID PYC1 YEAST
 AC P11154;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase 1 (EC 6.4.1.1) (Pyruvic carboxylase 1) (PCB 1).
 GN PYC1 OR PYV OR YGL062W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae;
 OX NCBI_TaxID=4932;

RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=88298805; PubMed=3042770;
 RA Lim F., Morris C.P., Ochiodoro F., Wallace J.C.;
 RT "Sequence and domain structure of Yeast pyruvate carboxylase."; J. Biol. Chem. 263:11493-11497(1988).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=S288C; PubMed=9234674;
 RC MEDLINE=97377993; PubMed=9234674;
 RX Feuerhann M., de Montigny J., Potier S., Souciet J.-L.;
 RA "The characterization of two new clusters of duplicated genes suggests a 'Lego' organization of the yeast *Saccharomyces cerevisiae* chromosomes."; Yeast 13:861-869(1997).
 RL [3]
 RN SEQUENCE OF 1003-1178 FROM N.A.
 RP MEDLINE=87241529; PubMed=3036136;
 RA Morris C.P., Lim F., Wallace J.C.;
 RT "Yeast pyruvate carboxylase: gene isolation."; Biochem. Biophys. Res. Commun. 145:390-396(1987).
 RL [4]
 CC -!- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction, involving the ATP-dependent carboxylation of the covalently attached biotin in the first step and the transfer of the carboxyl group to pyruvate in the second.
 CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3) (-) = ADP + phosphate + oxaloacetate.
 CC -!- COFACTOR: Biotin and zinc.
 CC -!- PATHWAY: Gluconeogenesis.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPONAMIDE TRANSFERASES AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J03889; AAA34843.1; -;
 DR EMBL; Z72584; CAA96765.1; -;
 DR PIR; S64066; QYBYP.
 DR HSP; P24182; IBNC.
 DR GenBank; U14110; -;
 DR SGD; S0003030; PYCL.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR GO; GO:0004736; F:pyruvate carboxylase activity; IDA.
 DR InterPro; IPR001882; Biotin_BS.
 DR InterPro; IPR005482; Biotin_Carb_C.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR005479; Cphase_L_D2.
 DR InterPro; IPR005481; HMGL-like.
 DR InterPro; IPR003379; PYC_OADA.
 DR InterPro; IPR005930; Pyruv_carbox.
 DR Pfam; PF02785; Biotin_Carb_C; 1.
 DR Pfam; PF00354; biotin_lipoyl; 1.
 DR Pfam; PF00289; Cphase_L_chain; 1.
 DR Pfam; PF02786; Cphase_L_D2; 1.
 DR Pfam; PF00682; HMGL-like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR TIGRfams; TIGR01235; pyruv carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPSASE_1; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 KW Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding; Zinc; Multigene family.
 FT NP_BIND 182 187 ATP (POTENTIAL).
 FT ACT_SITE 312 312 BY SIMILARITY.
 FT BINDING 1135 1135 BIOTIN (BY SIMILARITY).

FT CONFLICT 462 462 T -> G (IN REF. 1).
 FT CONFLICT 493 493 R -> D (IN REF. 1).
 FT CONFLICT 595 595 R -> A (IN REF. 1).
 FT CONFLICT 619 619 E -> Q (IN REF. 1).
 FT CONFLICT 664 664 G -> S (IN REF. 1).
 FT CONFLICT 772 772 A -> R (IN REF. 1).
 FT CONFLICT 879 879 E -> Q (IN REF. 1).
 FT CONFLICT 909 909 Q -> K (IN REF. 1).
 SQ SEQUENCE 1178 AA; 130099 MW; BC7110ASAFB23E04 CRC64;
 Query Match 42.4%; Score 2457; DB 1; Length 1178;
 Best Local Similarity 46.7%; Pred. No. 8.5e-124;
 Matches 541; Conservative 173; Mismatches 401; Indels 44; Gaps 16;
 QY 14 KILVANRGRIAVPAFAALETGAATVAIYPRDRSGSFHRSFASEAVRICTEG--SPVKAY 71
 DB 21 KILVANRGEIPRIIFRTAHELNSMTVAIYSHEDRLSTHKQKADAEVIVEVQYTPVGAY 80
 QY 72 LDIDEIIGAKKVKADAIYGYGLFSENLAQARECAENGITFIGPTPEVLDTGDKSRV 131
 DB 81 LAIDEIISTAQKHQVDFIHFGYGLFSENSEFADKVKVAGITWIGPPAEVIDSVGDKVSAR 140
 QY 132 TAAKAGLPVLAEST-PSKNIDIVKASBGQYPIFKAVAGGGRGMEFVASPDELRLK 190
 DB 141 NLAAKANVPTVPGTPIETVEEALDFVNEYGPVFIKAAFGGGGRMVRREGDDVADA 200
 QY 191 ATASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVLHYERDCSLQRHOKVYE 250
 DB 201 FORATSEARTAFNGTCFVERFLDKPKHIEVQLLDNHNHNVHLPERDCSVQRHOKVYE 260
 QY 251 IAPAQHLDPRLPRICADAVKFCRSIGYOGAGIVEELVDEKGNHVFIEKNPRIQVHTVT 310
 DB 261 VAPAKTLPREVRDAITLTDVAKLAKGCGYRNAGTAEFLVDNQNHYFIEINPRIQVHTIT 320
 QY 311 EEVTEVDLVKAQMLAAGATKELGTQDKIKTHGAALQCRITTEDPNNNGFRPDGTITA 370
 DB 321 EITGIDIVAQIQAAGASLPQLGLFQDKITTRGFAIQCRITTEDPAKNQFQDTGRIEV 380
 QY 371 YRPGGAGVLDGA-AQLGSEITAHFDSMLVMTCTGSDFFETAVARAQALAEFTVSGVA 429
 DB 381 YRSAGNGVELDGGNAYAGTIISPHYDSMLVKSCSGSTYEVVRRKMRALIEFRIGVK 440
 QY 430 TNIGFLRALIREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYLDVTVN---KP 485
 DB 441 TNIPFLTLTNPFIEGTWTTFIDDTQLFQWSSQNRAQLLHVLADVAVNGSSIKG 500
 QY 486 HGVPRKDVAAPIDKLPNIKDL-----PLPRGSRDLKQLGPAFAARDLREQDALA 535
 DB 501 QIGLPKLKSNP--SVPHLHDAQGNVNTKSPSGMRQVLLKGPFAEFARQVRQFNGTL 558
 QY 536 VTDTTPRDAHOSLLATRVRSFALKPAEAAVAKLTPELLSVAEWGGATYDVAMFLEDPW 595
 DB 559 LMDTWRDAHQSLLATRVTRHDLTATPTTAHALAGRFALCEWGGATFDVAMFLEDPW 618
 QY 596 DRDLREAMPNVIQMLRGRNTVGYTPYPSVCRFAFKVKAASSGVDIFRIFDALNDVS 655
 DB 619 ERLKRLASLVENIPFQMLLRGANGVAYSLPDAIDHFVKQAKDNGVDIFRVEDALNDLE 678
 QY 656 QMRPAIDAVLETNTAVAEVAMAYSGDLDNENKLYTLDDYVKMAEEIVKSGAHILAKDM 715
 DB 679 QLVGVDAVKAG-GVVEATVCFSGDMLQPKK-YLDDYILEIAEKIVQMGTHILGKDM 736
 QY 716 AGLRPAAVTKLVTALRREF-DLPVHVHTHTDAGGOLATYFAAAQAGADVDGASPLSG 774
 DB 737 AGTMKPAAKLLIGSLRAKYPDLPPIHVHTHDSAGTAVASMTACALAGADVVDVAINSMSG 796
 QY 775 TTSPSLSAIVAAFAHTRRTDGLSLAEVSDLEPYWEAVRGLYLPFESGFGPTGRVVRHE 834
 DB 797 LTSQPSINALASL-EGNIDTGINVRELDYAWEMRLLYSCFADLKGPDPEYQHE 855
 QY 835 IPGQSLNRAQATFALGLADRFEIIEDNYAAVNMELGRPTKVTSSKVGVDLHLHVGAG 894
 DB 856 IPGQSLTNLLFQAQQLGLGEQWAEKAYREANVLLGDIKVTPTSKVGVGLAQFVSNK 915

QY 895 VDPADFAADPOKIDIPDSVIAFLRGELNPPGQWPEPLRTRALEGRSEGKAPLTPVEE 954
 DB 916 LTSDDVRELANSLOFPDSDWDFEGLIGQPVGGFPFPRSDVLENK---RRKLCRGLE 972
 QY 955 QAHLDADDKSKERNLNRL-----LFPKPTTEPLEHRRRFGNTSALDDREFFY 1002
 DB 973 LEPTDLE--KIREDLQNRFGVDECDVASYNMYPRVYDFQKMEETYGDLISVLPTRSFLS 1030
 QY 1003 GLVEGRETLIRLPVTRTELLVRLDAISBPDK-QMRNVVANVGQIRPMVRDRSVESVT 1061
 DB 1031 PLETDDEEVVIEQGT-LTIKLAQVGLNKKTEREVYFDLNGEMRKIRVARSQKVT 1089
 QY 1062 ATAKKADSSNGHVAAPAGV-VTVVAGDEVKAGDAVAIIEMKQVEAITASVDGKID 1120
 DB 1090 VTKGKADHDLHIGAPMAGVIVEVKHGLIKKQGPVAVLSAMKVMIISSPSDQGVK 1149
 QY 1121 RVVPAATKVEGDLIVV 1139
 DB 1150 EVFVSDGENVDSDLLVLL 1168

RESULT 5

ID PYC2 YEAST STANDARD; PRT; 1180 AA.
 AC P32327.
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1994 (Rel. 30, last sequence update)
 DE Pyruvate carboxylase 2 (EC 6.4.1.1) (Pyruvate carboxylase 2) (PCB 2).
 GS PCB OR YBR218C OR YBR1507.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=S288c;
 RA Dubois E., el Bakoury M., Glansdorff N., Messenguy F., Pierard A.,
 RA Scherens B., Vierendeels F., EMBL/GenBank/DBSJ databases.
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBSJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RA Dubois E., el Bakoury M., Glansdorff N., Messenguy F., Pierard A.,
 RA Scherens B., Vierendeels F., EMBL/GenBank/DBSJ databases.
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBSJ databases.

QY MEDLINE=96128067; PubMed=8554526;
 RA Val D.L., Chapman-Smith A., Walker M.E., Cronan J.E. Jr.,
 RA Wallace J.C.;
 RT "Polymorphism of the yeast pyruvate carboxylase 2 gene and protein:
 RT effects on protein biotinylation.";
 RL Biochem. J. 312:817-825(1995).
 CC -!- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
 CC involving the ATP-dependent carboxylation of the covalently
 CC attached biotin in the first step and the transfer of the
 CC carboxyl group to pyruvate in the second.
 CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3) (-) = ADP + phosphate +
 CC oxaloacetate.
 CC -!- COFACTOR: Biotin and zinc.
 CC -!- PATHWAY: Gluconeogenesis.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- INDUCTION: By glucose.
 CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 CC AND CARBANYL PHOSPHATE SYNTHETASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X59890; CAA42544.1; -.
 DR EMBL; Z36087; CAA85182.1; -.
 DR EMBL; U35647; CAA49147.1; -.
 DR PIR; S46094; S46094.
 DR HSSP; P24182; LBNC.
 DR GeneOnline; 138751; -.
 DR SGD; S0000422; PYC2.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR GO; GO:0004736; F:pyruvate carboxylase activity; IDA.
 DR InterPro; IPR001882; Biotin BS.
 DR InterPro; IPR005482; Biotin carb C.
 DR InterPro; IPR000089; Biotin_lipoYL.
 DR InterPro; IPR005479; CPhase_L_D2.
 DR InterPro; IPR005481; CPhase_L_N.
 DR InterPro; IPR000891; HMGL-like.
 DR InterPro; IPR003379; PYC OADA.
 DR InterPro; IPR005930; Pyruv. carbox.
 DR Pfam; PF02785; Biotin carb C; 1.
 DR Pfam; PF0364; Biotin_lipoYL; 1.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR Pfam; PF00682; HMGL-like; 1.
 DR Pfam; PF02436; PYC OADA; 1.
 DR TIGRFAMs; TIGR01235; Pyruv. carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPSASE 1; 1.
 DR PROSITE; PS00867; CPSASE 2; 1.
 KW Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
 KW Zinc; Multigene family.
 FT NP_BIND 183 188 ATP (POTENTIAL).
 FT ACT_SITE 313 313 BY SIMILARITY.
 FT BINDING 1136 1136 BIOTIN (BY SIMILARITY).
 FT CONFLICT 15 15 S -> C (IN REF. 1).
 FT CONFLICT 132 132 D -> E (IN REF. 1).
 FT CONFLICT 238 238 N -> K (IN REF. 1).
 FT CONFLICT 268 268 L -> F (IN REF. 1).
 FT CONFLICT 546 546 S -> C (IN REF. 1).
 FT CONFLICT 642 642 N -> T (IN REF. 1).
 FT CONFLICT 771 773 GTA -> STR (IN REF. 1).
 FT CONFLICT 831 831 W -> R (IN REF. 1).
 FT CONFLICT 839 839 S -> P (IN REF. 1).
 FT CONFLICT 1001 1001 Y -> N (IN REF. 1).
 FT CONFLICT 1155 1155 K -> R (IN REF. 1).
 FT CONFLICT 1178 1178 Q -> P (IN REF. 1).
 FT CONFLICT 1180 1180 K -> KVIFTR (IN REF. 1).
 SQ SEQUENCE 1180 AA; 130166 MW; AD60DA3A60F5E001 CRC64;

Query Match 42.1%; Score 2436; DB 1; Length 1180;
 Best Local Similarity 45.5%; Pred No. 1,le-122;
 Matches 531; Conservative 183; Mismatches 408; Indels 44; Gaps 16;
 QY 7 STLPFKKILVANRGEIIVRAFAALETGAATVAIYPREDGSGFHRSEAFSEAVRIGTGG- 65
 DB 15 SLLEGKILVANRGEIPIRISAHLSMRTIATYSHEDRLSMRLKADEAYVIGEEGQ 74
 QY 66 -SPVXAYLIDIDIIIGAAKKVADATYPGVGFSENAQLARECAENGITFIGPTPEVLDT 124
 DB 75 YTPVGAYLAWDEIIEIAKKGVDFIHPGFGFSENSEFADKVVKAGIWIIGPPAEVIDSV 134
 QY 125 GDKSRVTAAKKAGLPLVLAEST-PSKNIDEIVKSAEGQTYPIFKVAVAGGGRGMRFVAS 183
 DB 135 GDKVSARHLAARANVTVPVTPGPIETVQEALDFVNEGYPIIKAFGGGGRGMRVRE 194
 QY 184 PHELKXLAETASREAAFGDGVVVERAVINPQHIQVILGDHTGEVVVHLYERDCSLQR 243
 DB 195 GDDVADAFORATSEARTAFNGTCFVERFLDKPKHIEVQLADNEGNNVHLLFERDCSVQR 254

244 RHQXVVEIAQHLDFELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFTEMNRI 303
 255 RHQXVVEIAQHLDFELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFTEMNRI 314
 304 QVEHTVTEVTEVDVKAQMLAAAGATLKLGLTQDKLTHGAALQCRITTEDPNNGFRP 363
 315 QVEHTVTEVTEVDVKAQMLAAAGATLKLGLTQDKLTHGAALQCRITTEDPNNGFRP 374
 364 DTGTTIAYRPGAGVRLDGA-AQIGETTAHFDMLNMTCRGSDFFETAVARAQALAE 422
 375 DTGRLEVYRSAGNGVRLDGNAYAGATISPHYSMLVKCSGSGTYEIVARRKIRALIE 434
 423 FTVSGVATNIGFLRALLREEDTSKFIATGFTADPHLLQAPPADDEQGRILDVLDVTV 482
 435 FRIRGVKTNIPHLLTLLTPVIEGTWTTFIDTLPQFQWSSQNRRAQKLLHYLDVAV 494
 483 N-KPHGVRRPKDVAAPIDKLPNIKDL-----PLPGSRDLKLGQPAAFARDL 528
 495 NGSSIKGQICLPLKLSNP--SVPHLHDAQGNVINTKSGAPSGMQVLEKGPSEFAQV 552
 529 RQDQALAVDTTFRDAHQSLATRVRSFALKPAEAVALKTPELLVSVEANGGATVDVNR 588
 553 RQNGTLLMDTTRDHAQSLATRVTHDLATIAPTTAAALAGAFALSCWGATFDVNR 612
 589 FLFEDPWRDLBELREAMPNNIQLLRGRNTVGYTPYDSCVCRATFVKEAASGVDIPRIF 648
 613 FLHEDPWRDLKLSLVNIPFQMLLRGANGVAYSSLPDNDIAIDHFVKQAKNGVDIPRVF 672
 649 DALNDVSNRPAIDAVLENTAVAEVAMAYSGDLPNEKLYTLDYLYKVAEEIVKSAH 708
 673 DALNDUEQKGVNAVAKKAG-GVREATVCYSDMLQPGKK-YNLDYLYLEVKEIVQMETH 730
 709 ILAIKDMAGLLRPAAVTKLVLTALREF-DLPVHVHTHTAGQLATYFAAAQAGADAVDG 767
 731 ILGIKDMAGTMRPAAKLLIGSLRTYRPLPIHVSHDSAGTAVASMTACALAGADVVDV 790
 768 ASAPLSGTTQSLSAIVAPFAHTRDTGLSLEAVSDLEPIWEAVRGYLLPFESTGPT 827
 791 AINSMSGLTQSLSINALLASL-EGNIDTGINVHVRELDAYWAEMRLLYSCFEADLKGP 849
 828 GRVYREIFPGQLSNLRAQATGALGRADREFELIEDNYAAVNEMLGRPTKVTSSKVVGDLA 887
 850 PEVYQHEIFGGQTLNLLFQAQQLGELGEQWAEKRAYREANYLLGDIKVVTPTSQVGDLA 909
 888 LHLVGAVDPADPAADPOKYDIPDSVIAFLRGLNPPGCGWPEPLRTRALGRSEKAPL 947
 910 QFMVSNKLTSDIRRLANSLDPDSVMDPFELIGQPYGPGPEPLRSVLRNK---RRKL 966
 948 TEVPEEQAHLDADDSKERRNSINRL-----LFPKTEEFLEHRRRFGNTGAL 995
 967 TCRPGLLEFPDL--KIBEDLQNRGDIQDQVSYNNYFVYEDFQKIBETYGDLVL 1024
 996 DREFFYGLVEGRETILRIPDVTPLVRLDAISEPDK-GMRNVVANVNGQIRPMVRD 1054
 1025 PTKNFLAPAEPPBEIEVTIEQGT-LIILQAVGLNKKTGQREYVYFELNGELKIRVAD 1083
 1055 RSVESVTATAEKADSNKGVHAAPFAGV-VTVVAGDEKVAAGDAVAIIEMKMEATITA 1113
 1084 KQONISVAKPADVHDTHQIGAPMAGVIEVVKHGLSVKKGESIAVLNKMMEVYSS 1143
 1114 SVDGKIDRVVPAATKVEGGDLIVV 1139
 1144 PADGQVKDVFIKGDSVDASDLLVL 1169

RESULT 6

PYC_PICPA STANDARD; PRT; 1189 AA.
 AC F78992;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Pyruvate carboxylase (EC 6.4.1.1) (Pyruvic carboxylase) (PCB).

GN PYC1.
 OS Pichia pastoris (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 CC NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98301182; PubMed=9639311;
 RA Merendez J., Delgado J., Gancedo C.;
 RT "Isolation of the Pichia pastoris PYC1 gene encoding pyruvate
 RT carboxylase and identification of a suppressor of the pyc
 RT phenotype";
 RL Yeast 14:647-654(1998).
 CC -1- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
 CC involving the ATP-dependent carboxylation of the covalently
 CC attached biotin in the first step and the transfer of the
 CC carboxyl group to pyruvate in the second (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: Biotin and zinc.
 CC -1- PATHWAY: Gluconeogenesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPONIC ACID TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y11106; CAA71993.1; -
 CC HSP: P24182; 1DV1.
 DR InterPro; IPR001882; Biotin_BS.
 DR InterPro; IPR005482; Biotin_Carb_C.
 DR InterPro; IPR000089; Biotin_LipoYL.
 DR InterPro; IPR005483; CPhase_L.
 DR InterPro; IPR005479; CPhase_L_D2.
 DR InterPro; IPR005481; CPhase_L_N.
 DR InterPro; IPR000891; HMGL-Like.
 DR InterPro; IPR003379; PYC_OADA.
 DR InterPro; IPR005930; Pyruv_carbox.
 DR Pfam; PF02785; Biotin_Carb_C; 1.
 DR Pfam; PF00364; biotin_LipoYL; 1.
 DR Pfam; PF00289; CPhase_L_chain; 1.
 DR Pfam; PF02786; CPhase_L_D2; 1.
 DR Pfam; PF00682; HMGL-Like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR PRINTS; PR00098; CPASASE.
 DR TIGRfams; TIGR01235; Pyruv_carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPASASE_1; 1.
 DR PROSITE; PS00867; CPASASE_2; 1.
 KW Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
 KW Zinc.
 FT NP_BIND 185 190 ATP (POTENTIAL).
 FT ACT_SITE 315 315 BY SIMILARITY.
 FT BINDING 1140 1140 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 1189 AA; 131400 MW; 8B6E858079657914 CRC64;
 Query Match 41.3%; Score 2391; DB 1; Length 1189;
 Best Local Similarity 45.3%; Pred. No. 2.9e-120;
 Matches 526; Conservative 177; Mismatches 423; Indels 34; Gaps 16;
 Qy 6 SSILPAPKILVANRGEIAVRAAPRAALETGAATVAIYPREDRGSFRRSEAVRGTSG 65
 Db 16 SSLLGTWNKILVANRGEIPIRIFRTAHELSMTVAIYSHEDRLSMRLKADGAYVGERG 75
 Qy 66 --SPVKAYIDIRIIGAARKVKADATYPGVFLSENAQALRECAENGITFIQPTPEVLDL 123
 Db 76 QYSPVQAYLAIDELIKIAVKHVNMIHPGVFCSENSEFARKVEENGILWVGPSTVIDA 135

DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR PRINTS; PR00098; CPSASE.
 DR TIGRFAMs; TIGR00514; accC; 1.
 DR PROSITE; PS00866; CPSASE.1; 1.
 DR PROSITE; PS00867; CPSASE.2; 1.
 DR Ligase; Multifunctional enzyme; Glucosyltransferase; Magnesium; Pyruvate;
 KW ATP-binding; Complete proteome.
 FT NP_BIND 162 167 ATP (POTENTIAL).
 FT ACT_SITE 291 291 POTENTIAL.
 SQ SEQUENCE 501 AA; 55402 MW; 04D2E401892F872F CRC64;

Query Match 18.6%; Score 1079; DB 1; Length 501;
 Best Local Similarity 48.2%; Pred. No. 1.3e-50;
 Matches 218; Conservative 82; Mismatches 146; Indels 6; Gaps 4;

QY 12 FKILVANRGEIAVRAFAALETGAATVAIYPRDGRGSHRFSFAEAVRIGTEGSPVKAY 71
 Db 2 FNKVLIANRGEIALRILRACHEMGATIAVSTVDNALHVLQDAEAVCIG-EPASAKSY 60
 QY 72 LDIDEITGAACKVKADAIYPGYGLFSENAQLARCAENGTITFTGPTPEVLDLTGDKSRV 131
 Db 61 LNTDAILNVAEKAKVDALHFGYGLAENAEFAFAVKKAGFEFIGNPDIAEAMGSKINAK 120
 QY 132 TAACKAGLPVLAESTPS-KNIDIVKSAEQTYPIFKAVAGGGRGMRFRVSPDELRLK 190
 Db 121 KIMKAGVPLIPSGEAGIEDIDAEIAEAIAGFPVVVKASAGGGMGMSVAYSKEELKEV 180
 QY 191 ATASRAEAAFGDGVAVYERAVINPQHIEVOITLGDHTGEVVLHYERDCSLQRHKKVVE 250
 Db 181 IEGARNIAKSAFGDPTVFEIKYLENPRHIEIQLLGKNGNIHLGDRCSIQRRHKKLIE 240
 QY 251 IAPAQHLDPRLDRICADAVKFCRSIYOGAGTVEFLVDEKGNHVFIEWNPRIQVEHTVT 310
 Db 241 EAFSPINTEELRGEAMIAKAGKAINYSAGTVEFLY-ENGNYFLENTRIQVEHTVT 299
 QY 311 EIVTEVDLVAQMLAAGATLKLGLTQDKIKTHGAALQCRITDPNNGFRPDTGTTA 370
 Db 300 EQVTGIDLKAMTKIAAG---EELTLKQEDVKIRGHAIECRINAEPLNDLFPVCGKIKL 356
 QY 371 YRSPGGAGVLDGAAQLGGITAHFDSMLVKMTCGSDFEETAVARAQALAEFTVSGVAT 430
 Db 357 YRSPGGGVLDGSGVGGAEIPYYDSMAKLITYGNSREAIARKKALREIVTIGVKT 416
 QY 431 NIGFLRALLREEDFTSKRIATGFIADPHLLQ 462
 Db 417 NIPFRAVLEENFLKGNISTHYVEQNMLKL 448

RESULT 8
 ID ACCC_ANASP STANDARD; PRT; 447 AA.
 AC Q06862;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
 DE carboxylase) (EC 6.4.1.2) (ACC).
 GN ACCC OR ALR0939.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93352435; PubMed=8102363;
 RA Gornicki P., Scappino L.A., Hasselkorn R.;
 RT "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
 RT sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
 RT protein.";
 RL J. Bacteriol. 175:5268-5272 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;

RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:203-213 (2001).
 CC -1- FUNCTION: This protein is a component of the acetyl coenzyme A
 CC carboxylase complex; first, biotin carboxylase catalyzes the
 CC carboxylation of the carrier protein and then the transcarboxylase
 CC transfers the carboxyl group to form malonyl-CoA.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -1- SUBUNIT: ACETYL-CoA CARBOXYLASE IS AN HETEROHOMER OF BIOTIN
 CC OF CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
 CC -1- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
 CC PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L14862; AAB51770.1; --
 CC EMBL; AP003584; BAB72896.1; --
 CC FIR; A53311; A53311.
 CC FIR; AH1923; AH1923.
 CC HSP; P24182; LBNC.
 CC InterPro; IPR004549; AccC.
 CC InterPro; IPR005482; Biotin carb. C.
 CC InterPro; IPR005479; CPase L D2.
 CC InterPro; IPR005481; CPase L N.
 CC Pfam; PF02785; Biotin carb. C. 1.
 CC Pfam; PF02789; CPSase_L_chain; 1.
 CC Pfam; PF02786; CPSase_L_D2; 1.
 CC TIGRFAMs; TIGR00514; accC; 1.
 CC PROSITE; PS00866; CPSASE.1; 1.
 CC PROSITE; PS00867; CPSASE.2; 1.
 KW Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
 KW Complete proteome.
 FT NP_BIND 163 168 ATP (BY SIMILARITY).
 FT ACT_SITE 293 293 BY SIMILARITY.
 FT AC_SITE 447 447 MW; 84541B3839E00F9 CRC64;
 SQ SEQUENCE 447 AA; 49104 MW; 84541B3839E00F9 CRC64;

Query Match 17.9%; Score 1036.5; DB 1; Length 447;
 Best Local Similarity 48.4%; Pred. No. 2.1e-48;
 Matches 216; Conservative 73; Mismatches 152; Indels 5; Gaps 3;

QY 12 FKILVANRGEIAVRAFAALETGAATVAIYPRDGRGSHRFSFAEAVRIGTEGSPVKAY 71
 Db 3 FDKILLIANRGEIALRILRACHEMGATIAVSTVDNALHVLQDAEAVCIG-EPASAKSY 61
 QY 72 LDIDEITGAACKVKADAIYPGYGLFSENAQLARCAENGTITFTGPTPEVLDLTGDKSRV 131
 Db 62 LNIPIIIAALTRNASAIHPGYGLFSENAKFAETCADHIIAFITGPTPEAIRLMDGKSTAK 121
 QY 132 TAACKAGLPVLAESTPSKNID-EIVKSAEGQTYPIFKAVAGGGRGMRFRVSPDELRLK 190
 Db 122 ETMQAGVPTVPGSEGLVETQEGLAKDI-GYPVMIKATAGGGRGMRFRVSPDEFVKL 181
 QY 191 ATEASREAAAGDGVAVYERAVINPQHIEVOITLGDHTGEVVLHYERDCSLQRHKKVVE 250
 Db 182 FLAAGEAGAAFGVAGVYIEKFIERPHIEFQIADNNGVNIHLGERDCSIQRNOKLLE 241
 QY 251 IAPAQHLDPRLDRICADAVKFCRSIYOGAGTVEFLVDEKGNHVFIEWNPRIQVEHTVT 310
 Db 242 EAPSPALDSDLREKMGQAANKAAQINVTGATIEFLDRLSGQYFWMNTRIQVEHPVT 301

DR	InterPro: IPR005479; CPase_L_D2.	
DR	InterPro: IPR005481; CPase_L_N.	
DR	Pfam: PF02785; Biotin_carb_C; 1	
DR	Pfam: PF02889; CPasease_1_chain; 1.	
DR	Pfam: PF02786; CPasease_L_D2; 1.	
DR	TIGRFAMs: TIGR00514; accC; 1.	
DR	PROSITE: PS00866; CPASE_1; 1.	
DR	PROSITE: PS00867; CPASE_2; 1.	
KW	Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;	
KW	ATP-binding; Complete proteome.	
FT	NP_BIND 161 166	ATP (POTENTIAL).
FT	ACT_SITE 290 290	POTENTIAL.
FT	SEQUENCE 506 AA; 57438 MW; 39B765F319235AD1 CRC64;	
DR	SEQUENCE	

Query Match	17.6%	Score 1019;	DB 1;	Length 506;
Best Local Similarity	47.9%;	Pred. No. 2	1e-47;	
Matches 213;	Conservative 73;	Mismatches 153;	Indels 6;	Gaps
Qy	12	FKKILVANRGEIAVRAFRAALETGAATVAIYPRDRGSPHRSPFASAEAVRIGTEGGSPVKAY		
Db	2	FSKILVANRGEIAVRVMRACRELGIKTGVGSSADKRAFRHYVADCCYIG-KADPRDSY		
Qy	72	LDTDEILIGAKKVKADAIYGGYGFSENAQLARACAENGITFIGPTPEVLDTGDKSRRAV		
Db	61	LNIDRIEIVAKKSGGAAIHFGYGFLEAENAPRCEEGIVFIGSPVEVIRTAGSKVRGR		
Qy	132	TAAKKAGLPVLAESTPSKNIDIVKSAEGCTYPIFVKAVAGGGGRMRFPVAGDELRKUAL		
Db	121	ESMQRAGVPVIPGSPKIDTVDEAKEMAEIKIGYPVAVKASGGGGIGIVVNSQEELEEAR		
Qy	192	TEASRAEAAFGDGAIVYVERAVINPQHIIEVOILGDHTGEVHLYPRDSCLSQRHKKVFE		
Db	181	RKSKLGESYFKDSTVYLEKYIARPHSIEVQILADQHGNNVHLGERECISIQRRHQKLIBE		
Qy	252	APAQHLDPEDLRDRI CADAVKFCRSIGYQGAGTVEFLVDEKGNHVFTEMNPRIOQVETVTE		
Db	241	APSPALNEEMREELGKLVAKGAREIGYTNAGITEFLY-ENGNYFYLEINSLRQLVEHTIT		
Qy	312	EVEVDLVKQAPMLAAGATLKEIGLGTQDKIKTHGAALQCRITTEDPNNQRPDPTGIITAY		
Db	300	VVTGIDIVYQIRIAYG--EELRHQGEDVAIRGHAIECRINAEDPVN-FYPSRSGRIILHY		
Qy	372	RSPGGAGVRLDGAQQLGGEITAHDFDSMLVKWTCRGSDPETAVARAQRALAEFTVSGVATV		
Db	356	RSPGGGIRVDSGIHMGYRIPPEYDSNISKLTAYGETREEARMKREALYEYIIEGVETN		
Qy	432	IGFLRALLREEDFTSKRIATGFIAD 456		
Db	416	IPHFVAVLNDEEFVRGNIIHTKFVEE 440		

RESULT 10	
PYCA_MEITH	
ID PYCA_METHTH STANDARD; PRT; 491 AA.	
AC O27939;	
AD 30-MAY-2000 (Rel. 39, Created)	
DT 30-MAY-2000 (Rel. 39, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
DS Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).	
GN PYCA OR MTH1917.	
CN Methanobacterium thermoautotrophicum.	
OS Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;	
OC Methanobacteriaceae; Methanothermobacter.	
CC NCBI_TaxID=187420;	
EN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=Delta H;	
RX MEDLINE=38037514; PubMed=9371463;	
RA Smith D.R., Doucetce-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.	
RA Alredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,	
RA Harricson D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,	
RA Safadara R., Vicare R., Wang Y., Wierzbowskij J., Gibson R.,	
RA Jiwanani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,	

RA McDougall S., Shiner G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics."; J.
J. Bacteriol. 179:7135-7155(1997).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-17.
RC STRAIN=Delta H;
RX MEDLINE=98148063; PubMed=9478969;
RA Mukhopadhyay B., Stoddard S.F., Wolfe R.S.;
RT "Purification, regulation, and molecular and biochemical
RT characterization of pyruvate carboxylase from Methanobacterium
thermoautotrophicum strain deltaH"; J. Biol. Chem. 273:5155-5165(1998).
RL J. Biol. Chem. 273:5155-5165(1998).
CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC GROUP TO PYRUVATE IN THE SECOND. THE MAXIMUM ACTIVITY IS AT PH 8
CC AND 60 DEGREES CELSIUS.
CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC oxaloacetate.
CC -1- COFACTOR: ATP, magnesium (or manganese or cobalt), pyruvate and
CC bicarbonate.
CC -1- ENZYME REGULATION: INHIBITED BY ADP AND ALPHA-KETOGLUTARATE.
CC -1- PATHWAY: Gluconeogenesis.
CC -1- SUBUNIT: Heterooctamer of four A and four B subunits.
CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC AND CARBAMYL PHOSPHATE SYNTHETASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; A800942; AAB86377.1; -
DR HSP; P24182; IBCN.
DR InterPro: IPR004549; AccC.
DR InterPro: IPR005482; Biotin carb C.
DR InterPro: IPR005479; CPase_L D2.
DR InterPro: IPR005481; CPase_L N.
DR Pfam: PF02785; Biotin carb C; 1.
DR Pfam: PF00289; CPase_L chain; 1.
DR Pfam: PF02786; CPase_L D2; 1.
DR TIGRfam: TIGR00514; accC; 1.
DR PROSITE: PS00866; CPASE 1; 1.
DR PROSITE: PS00867; CPASE 2; 1.
KW Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
KW ATP-binding; Complete proteome.
FT NP_BIND 162 167 ATP (BY SIMILARITY).
FT ACT_SITE 291 291 POTENTIAL.
SQ SEQUENCE 491 AA; 54656 MW; 5789C34DA7475C2E CRC64;

Query Match 17.2%; Score 997; DB 1; Length 491;
Best Local Similarity 45.9%; Pred. No. 3.1e-46;
Matches 205; Conservative 79; Mismatches 157; Indels 6; Gaps 4;

QY 12 FKILVANRGEIAVRAAFRAALETGAATVAIYPRDRGSRFSRFASEAVRIGTEGSPVKAY 71
DB 2 FSKILVANRGEIAVRAACRELGIKSVAVYSEADKXALFTRYADEAYEIG-KPAPSQSY 60

QY 72 LDDEITIGAKVKADAIYGVGFSENQALRECAENGITIGTPEVLDLTGDKSRV 131
DB 61 LRIDRLVAVKAGABAIHFGVGLAENPLRGECEKQKILGPKGSIEMAGDKITSK 120

QY 132 TAAKAGLFLVAESTPS-KNIDBIKVSABGQTYPIFKVAVAGGGGGRMRFVSPDELRL 190
DB 121 KLMKAGVPIPTDQGVDPDEAARIADISGVPIIKASAGGGGIGMRAVVEDELIRA 180

QY 191 ATEASREAAFGDGAIVYVERAVINPQHVLEIQLDGHTEGVHLVERDCSLQRKQVVE 250

Db 181 MESTQVSAASAFDPTFYIEKYLPRHIEFQVMADESGNVHLADRECSIORRHQKLE 240
QY 251 IAPAQHLDPRLDRICADAVKFCRSIGYQAGVVEFLVDEKGNHVFIEKNPRIQVEHTVT 310
Db 241 EAPSPIMTPELRERMGSAVKAAYIGYENACTVEFLY-SNGDFYFLENNTRIQVEHPT 299
QY 311 EEVTEVDLVKQMLAAGATLKLGLTQDKIKTHGAALOCRTITTEDPNNGPRDPTGTTA 370
Db 300 EVITGVDLVKEQIRVASG---ELRFTQKDINIRGHAIECRINAENPLADFAFPKGTG 356
QY 371 YRSPGAGVRLDGAALGCEITAHFDFMLVMTKGTGSDPETAVARAQALASFTVSGVAT 430
Db 357 YRSPGGIGVYRVDGVTMYEIPFYDSMISKLIVMGMDRQEAINEWKRLSEVILGVKT 416
QY 431 NIGFLRALLREEDFTSKRIATGTFIADH 457
Db 417 TIPFKAIWNEAFRRGELHFTHFVDEY 443

RESULT 11
ACCC_HAEIN
ID ACCC_HAEIN STANDARD; PRT; 448 AA.
AC P43873;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE carboxylase) (EC 6.4.1.2) (ACC).
GN ACC OR H10972.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Midan J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.D., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd".
RL Science 269:496-512(1995).
CC -1- FUNCTION: This protein is a component of the acetyl coenzyme A
CC carboxylase complex; first, biotin carboxylase catalyzes the
CC carboxylation of the carrier protein and then the transcarboxylase
CC transfers the carboxyl group to form malonyl-CoA (by similarity).
CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC -1- SUBUNIT: ACETYL-CoA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (BY SIMILARITY).
CC -1- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC PHOSPHATE SYNTHETASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32778; AAC22632.1; -
DR PIR; P64105; P64105.

```
DR HSP; P24182; 1BNC.
DR TIGR; H10972; -
DR InterPro; IPR004549; AccC.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR Pfam; PF02785; Biotin_carb_C_1.
DR Pfam; PF02889; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR TIGRPFAMs; TIGR00534; accC; 1.
DR PROSITE; PS00866; CPASE_1; 1.
DR PROSITE; PS00867; CPASE_2; 1.
KW Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
Complete proteome.
FT NP_BIND 163 168 ATP (POTENTIAL).
FT ACT_SITE 292 292 POTENTIAL.
SQ SEQUENCE 448 AA; 49108 MW; 28497E2A31ED96D1 CRC64;

Query Match 16.3%; Score 946; DB 1; Length 448;
Best Local Similarity 47.6%; Pred. No. 1.4e-43;
Matches 208; Conservative 61; Mismatches 154; Indels 14; Gaps 7;

QY 13 KTLVANRGRIAVRAFAALETGAATVAIYPREDRGSPHRSFASAEAVRIGTSGSPKAYL 72
DB 3 EKVVIANRGRIARIRACKELGIKVAVHSTADRLKHVLLADETICIGPAPS-AKSYL 61

QY 73 DIDEITGAKKVKADAIYPGYFLSNAQLARECAENGITFIGPTPEVLDTGDKSRVAV 132
DB 62 NIPAITAAAEVTCADAIHPGYFLSNAQDPFAEQVRSRGTFIGPTADVIRLMDKVSATK 121

QY 133 AAKAGLPLV--AESPTSKNIDIVKSAEGQTPVIFVKAVAGGGGGMRFVSPBELRLK 190
DB 122 AMKAGVPCVPGSDGVSNDIAKNTAKRIGVPIIIRKASGGGGGGMRFVSEDALES 181

QY 191 ATEASREAAFGDGVAVYERAVINPQHIEVQLGDHTGEVHLYERDCSLQRHKKVYE 250
DB 182 IAMTKAEAKAFAFNMDVMYMEKYLENPRHVEIQVLADTHGNVYLAERDCSMQRHKKVYE 241

QY 251 IAPQHLDPRLRI---CADAVKFCISGYCAGTVEFLVDEKGNHVIENPRLQVHEH 307
DB 242 EAPAPGITEVRRDIDGSRANA---CVEIGYRAGTFFELY-ENGFEFFIEMNTRIQVHEH 297

QY 308 TVTEEVTEVLDVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEPNNNGFRPDGT 367
DB 298 PVTEMITGVLDVKEQLRIAGL---PISFKQEDIKVKGHAMECRINAEDPKT-FLFSPGK 353

QY 368 ITAYRSPGGAGVRLDGAQLGGIHTAHPDSMLVWMTCRGSDPETAVARAQRALAEFTVSG 427
DB 354 VNHLSFGGLGVWRDWSHVYGGYTPPHYDSMIAKLITYGDTREVAIRRMQNALSETIIDG 413

QY 428 VATNIGFLRALREEDF 444
DB 414 IKTNIPHEILLEDENF 430

RESULT 12
ACCC_ECOLI STANDARD; PRT; 449 AA.
AC P24182;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE carboxylase) (EC 6.4.1.2) (ACC).
GN ACCC OR FABG OR B3256.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=K12;
RX MEDLINE=92052166; PubMed=1682920;
```

```
RA Kondo H., Shiratsuchi K., Yoshimoto T., Masuda T., Kitazono A.,
RA Teuru D., Anai M., Sekiguchi M., Tanabe T.;
RT "Acetyl-CoA carboxylase from Escherichia coli: gene organization and
RT nucleotide sequence of the biotin carboxylase subunit.",
RL Proc. Natl. Acad. Sci. U.S.A. 88:9730-9733(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112819; PubMed=1370469;
RA Li S.-J., Cronan J.E. Jr.;
RT "The gene encoding the biotin carboxylase subunit of Escherichia coli
RT acetyl-CoA carboxylase.",
RL J. Biol. Chem. 267:855-863(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Best E.A., Knauf V.C.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Ferna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA May B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.",
RL Science 277:1453-1474(1997).
RN [5]
RP SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.",
RL Electrophoresis 18:1259-1313(1997).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=94347758; PubMed=7915138;
RA Waldrop G.L., Raymont I., Holden H.M.;
RT "Three-dimensional structure of the biotin carboxylase subunit of
RT acetyl-CoA carboxylase.",
RL Biochemistry 33:10249-10256(1994).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=20283656; PubMed=10821865;
RA Thoden J.B., Blanchard C.Z., Holden H.M., Waldrop G.L.;
RT "Movement of the biotin carboxylase B-domain as a result of ATP
RT binding.",
RL J. Biol. Chem. 275:16183-16190(2000).
CC -|- FUNCTION: This protein is a component of the acetyl coenzyme A
CC carboxylase complex; first, biotin carboxylase catalyzes the
CC carboxylation of the carrier protein and then the transcarboxylase
CC transfers the carboxyl group to form malonyl-CoA.
CC -|- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -|- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC -|- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROMER OF BIOTIN
CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
CC -|- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M79446; AAA23748.1; -.
CC EMBL; M80458; AAA23409.1; -.
CC EMBL; M83198; AAA23746.1; -.
CC EMBL; U18997; AAA58059.1; -.
DR
DR
DR
DR
```

```

DR EMBL; AB000404; AAC76288.1; -.
DR PIR; JS0632; JS0632.
DR PDB; 1BNC; 30-AUG-95.
DR PDB; 1DV1; 09-JUN-00.
DR PDB; 1DV2; 09-JUN-00.
DR PDB; 1K69; 05-DEC-01.
DR EcGene; EGI0276; accC.
DR InterPro; IPR004549; AccC.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF02789; CPase_L_N.
DR Pfam; PF02786; CPase_L_D2; 1.
DR TIGRfam; TIGR00514; accC; 1.
DR PROSITE; PS00866; CPASE_1; 1.
DR PROSITE; PS00867; CPASE_2; 1.
DR Fatty acid biosynthesis; Ligase; Biotin; ATP-binding; 3D-structure;
KW Complete proteome.
FT NP_BIND 163 168 ATP (POTENTIAL).
FT ACT_SITE 222 222 POTENTIAL.
FT CONFLICT 260 261 CA -> SR (IN REF. 2).
FT CONFLICT 313 313 L -> M (IN REF. 1).
FT STRAND 4 7
FT HELIX 11 24
FT TURN 25 25
FT STRAND 27 33
FT HELIX 34 36
FT TURN 37 38
FT HELIX 40 44
FT STRAND 47 52
FT HELIX 56 58
FT TURN 59 61
FT HELIX 63 73
FT TURN 74 74
FT STRAND 77 79
FT HELIX 84 87
FT TURN 89 97
FT TURN 98 99
FT STRAND 101 102
FT HELIX 107 114
FT HELIX 116 125
FT TURN 126 127
FT STRAND 130 130
FT STRAND 133 133
FT HELIX 142 152
FT STRAND 156 160
FT STRAND 172 172
FT HELIX 175 188
FT STRAND 200 202
FT STRAND 208 217
FT STRAND 222 234
FT TURN 235 236
FT STRAND 237 243
FT TURN 246 247
FT HELIX 250 267
FT TURN 268 268
FT STRAND 270 280
FT TURN 281 282
FT STRAND 283 290
FT TURN 295 296
FT HELIX 297 304
FT HELIX 308 316
FT TURN 317 318
FT HELIX 325 327
FT STRAND 333 340
FT STRAND 342 342
FT TURN 344 346
FT STRAND 349 349
FT STRAND 352 353
FT STRAND 356 358
FT TURN 363 364
FT STRAND 365 365

```

Query Match 16.2%; Score 938; DB 1; Length 449;
Best Local Similarity 46.2%; Pred. No. 3.9e-43;
Matches 200; Conservative 69; Mismatches 156; Indels 8; Gaps 5;

14 KILVANRGEIARAFRAALETGAATVAIYPRDRSGSFHRSFASAEVRIIGTEGSPVKAYLD 73
4 KIVIANRGEIARILIRACKELGIKTAVHSSADRLKHVLLADETVICIGPAPS-VKSYLN 62
74 IDBITGAAKVKADAIYPGYGLSENALARECAENGITFTGPTPEVLDLTGDKSRVTA 133
63 IPAILISAAEITGAVAIHPGYGLSENANFAEQVRSQFIFIGFKAETIRLMGDKVSAIAA 122
134 AKKAGLPVL--AESTPSKNIDEIVKSAEQTYPIFVKAVAGGGGRGMRFPVSPDELRLKLA 191
123 MKKAGVPCVPGSDGGLDDMDKNRAIKRIGYFVVIKASGGGGGRGMRVVRGDAELAQSI 182
192 TEASREAAAPGDGAVYVERAVINPQHIEVQLDHTGTVHLYYERDSCSLQRHOKVVEI 251
183 SMTRAERAKAFAFSDVMYMEKYLENPRHVEIQVLADQGNAYILAERDCSMQRHOKVVEE 242
252 APACHLDPELRDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFIEHNPRIQVHTVTE 311
243 APAGITPELRRIYIGERCACAVDIGRGAGTFELF-ENGFEFYIEMWTRIQVHEPVT 301
312 EVTEVDLVKAMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNNGFRPDTGTITAY 371
302 MITGVDLIKEQLRIIAAG--QPLSIKQBEHVVRGHAVERCINAEDPNT-FLFSPGKITRF 357
372 RSPGGAGVRLDGAALGGGEITAHFDSMLVKMFCRSDDEFATAVAAQALAEFTVSGVAVN 431
358 HAPGGFVWRWESHVYAGTYVPFYIDSMIGKLCYGENRDVATARMKNALQELIIDGIKTN 417
432 IGFRLALLREEDF 444
418 VDLQIRIMDENF 430

RESULT 13
ACCC_ECO57 STANDARD; PRT; 449 AA.
AC Q8X9B6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE carboxylase) (EC 6.4.1.2) (ACC).
GN ACCC OR Z4616 OR ECS4128.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:523-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=O157:H7 / RIMD 0509952;
 RC MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.,
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -!- FUNCTION: This protein is a component of the acetyl coenzyme A
 CC carboxylase complex; first, biotin carboxylase catalyzes the
 CC carboxylation of the carrier protein and then the transcarboxylase
 CC transfers the carboxyl group to form malonyl-CoA.
 CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
 CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
 CC OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (By similarity).
 CC -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
 CC PHOSPHATE SYNTHETASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; AE005553; BAG58384.1; -;
 DR EMBL; AP002564; BAB3551.1; -;
 DR PIR; H91144; H91144.
 DR InterPro; IPR004549; AccC.
 DR InterPro; IPR005482; Biotin carb C.
 DR InterPro; IPR005479; Cbase_L_D2.
 DR InterPro; IPR005481; Cbase_L_N.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF02786; Cpsase_L_chain; 1.
 DR Pfam; PF02786; Cpsase_L_D2; 1.
 DR TIGRFAMs; TIGR00514; accC; 1.
 DR PROSITE; PS00866; CPSASE_1; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 KW Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
 KW Complete proteome.
 FT NP_BIND 163 168 ATP (POTENTIAL).
 FT ACT_SITE 292 292 POTENTIAL.
 SQ SEQUENCE 449 AA; 49337 MW; BC5716323F6233ED CRC64;
 Query Match 15.28; Score 936; DB 1; Length 449;
 Best Local Similarity 46.23; Pred. No. 5e-43;
 Matches 200; Conservative 69; Mismatches 156; Indels 8; Gaps 5;
 QY 14 KILVANGETAVRAFAAETGAATVAIYPREDRSGFHRSEAVRIGTEGSPKAYLD 73
 DB 4 KIVIANGETALRLTRACKELGKTKVAVHSSADRLKVLADETVCIGPAPS-VKSYLN 62
 QY 74 IDEIIGAARKVKADAIYPGVGFSENQALRECAENGITPTPEVLDLTDGKSAVTA 133
 DB 63 IPAISAAETGVAIHPGVGFSENANFAEQVERSGFIFGPKAETIRLMGDKVSAIAA 122
 QY 134 AKKAGLPVL--ASTPSKNIDEIVKSAEGQTFIFVKAVAGGGGRGVRFVASPDELKLA 191
 DB 123 MKKAGVPCVSGDGLGDDMKRAIAKRGYPVIKASGGGGGRGVVRGDAELAQSI 182
 QY 192 TEASREAEAFGGCAVTVERRAVINPQHIEVQILGDHTGEVHLHYERDCSLQRHKKVET 251
 DB 183 SMTRAEAKAAAFNDVMVYMEKYLENPRHVEIQVLADGGNSIYLAERDCSNQRHKKVVE 242

QY 252 APAQHLDPELRDRICADAVKCRSTCYOGACTVEFLVDEKGNHVFIEKNPRIQVHEHTYE 311
 DB 243 APAPGITPELRRIYIGERCACACVDIGYRGAGTFEFLF-ENGEFYFIEKNTRIQVHEHTYE 301
 QY 312 EYTEVDLVKAQMLAAGATLKEIGLTQDKIKTHGAALQCRITTEDPNNNGFRPDTGTTITAY 371
 DB 302 MITGVDLIKEQLRIAG--QPLSIKQEVHVRGHAVECRINAEDPNT-FLPSPGKIITRF 357
 QY 372 RSPGGNGVRLDGAALGGEITAHFDSMLVMTKRCSDFTETAVARQALAEFTVSGVATN 431
 DB 358 HAPGGFGVRWESHIVAGTVPPYDYMIGKLCYGENRDVAIARMKNALQELIIDGKTN 417
 QY 432 IGFLRALLREEDF 444
 DB 418 VDLQIRIMNDENF 430
 RESULT 14
 ACCC PSBAE
 ID ACCC PSBAE STANDARD; PRT; 449 AA.
 AC P37798;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
 DE carboxylase) (EC 6.4.1.2) (ACC).
 OS Pseudomonas aeruginosa.
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OS Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAOI;
 RX MEDLINE=94042851; PubMed=7693652;
 RA Best E.A., Knauf V.C.;
 RT "Organization and nucleotide sequences of the genes encoding the
 RT biotin carboxyl carrier protein and biotin carboxylase protein of
 RT Pseudomonas aeruginosa acetyl coenzyme A carboxylase";
 RL J. Bacteriol. 175:6881-6889(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAOI;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Hickey C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
 RA Garber R.L., Brinkman F.S.D., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RA "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
 RA opportunistic pathogen";
 RL Nature 406:959-964(2000).
 CC -!- FUNCTION: This protein is a component of the acetyl coenzyme A
 CC carboxylase complex; first, biotin carboxylase catalyzes the
 CC carboxylation of the carrier protein and then the transcarboxylase
 CC transfers the carboxyl group to form malonyl-CoA.
 CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
 CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
 CC OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
 CC -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
 CC PHOSPHATE SYNTHETASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----

or send an email to license@ib-sib.ch).

```

CC  -----
CC  EMBL; L14612; AAA16041.1; .
DR  EMBL; AE004898; AAG08233.1; -.
DR  PIR; B49342; B49342.
DR  HSSP; 224182; 1BNC.
DR  InterPro; IPR004549; AccC.
DR  InterPro; IPR005482; Biotin carb C.
DR  InterPro; IPR005479; CPase_L_D.
DR  InterPro; IPR005481; CPase_L_N.
DR  Pfam; PF02785; Biotin carb C; 1.
DR  Pfam; PF00289; CPase_L chain; 1.
DR  Pfam; PF02786; CPase_L_D2; 1.
DR  TIGRfams; TIGR00514; AccC; 1.
DR  PROSITE; PS00866; CPASE_1; 1.
DR  PROSITE; PS00867; CPASE_2; 1.
KW  Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW  Complete proteome.
FT  NP_BIND 163 168 ATP (POTENTIAL).
FT  ACT_SITE 292 292 POTENTIAL.
SQ  SEQUENCE 449 AA; 48888 MW; 3B04C77785C73541 CRC64;
Query Match 16.0%; Score 924; DB 1; Length 449;
Best Local Similarity 45.9%; Pred. No. 2.2e-42;
Matches 199; Conservative 67; Mismatches 160; Indels 8; Gaps 5;
QY 13 KKILVANRGEIAVRAALETGAATVALYPRDRGSPHRSFAGEAVRIGTEGSPVKAYL 72
DB 3 EKVLNARGEIALRLIRACKELGIKTVAHVSTADRELHSLADESVICG-PAPATQSYL 61
QY 73 DIDEIIGAKKVADALVPGYGLFSENQAQARECAENGIITIGTPEVLDLTGDKSAVT 132
DB 62 QIPALIAAAEVGTATAIHPGYGLFAENADFAEQIERSGTFVGTAEVIRLMGDKVSAKD 121
QY 133 AAKKAGLPLVARSTESKNIDE--IVKSAEGQYPIFFKAVAGGGGRGRRFVASPDELRL 190
DB 122 AMKRGVPIVPSDGLPDEDETALAREVCYPIVIAAGGGGGRGNRVYVDSSELKS 181
QY 191 ATEAREAEAAFGDGAUVYVERAVINPOHIEVOILGDHTGEVVHLYVERDCSLQRHQKVE 250
DB 182 AKLIRTEAGAAFGNPNVLEKLTNPFRHVEVQVLSGQGNAILHGLDRDCSLQRHQKVE 241
QY 251 IAPACHLDELDRICADAVKCRSIGVGGAGTVEFLVDEKGNHVFIMNPRIOVEHTVT 310
DB 242 EAPAGCIDKAQEVFARVCVACIIGVAGAGTFEFLY-ENGRFFIENRVRQVEHPVS 300
QY 311 EEVTEYDLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNFRPDTGITTA 370
DB 301 ENVTGVDIVKEMLRISAG--EKLSIQEDVVIRGHAELECRINAEDPKT-FMPSPGKVKH 356
QY 371 YRSPGAGYRLDGAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 430
DB 357 FHAPGNGVRVDSHLYSGYSVPNPVDSLVGKVITYGADRDALARNALDELIVDGIKT 416
QY 431 NIGFLEALLREEDF 444
DB 417 NTELHKLDVDAAF 430

```

RESULT 15

```

MCCA_HUMAN
ID MCCA_HUMAN STANDARD; PRT; 725 AA.
AC Q96R03; Q9H959; Q9NS97;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Methylcrotonyl-CoA carboxylase alpha chain, mitochondrial precursor
DE (EC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 1) (MCCase alpha
DE subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
GN MCCC1 OR MCCA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS MCGI ARG-325 AND SER-385.
RX PubMed=1170888;
RA Gallardo M.E., Desviat L.R., Rodriguez J.M., Esparza-Gordillo J.,
RA Perez-Cerda C., Perez B., Rodriguez-Pombo P., Criado O., Sanz R.,
RA Morton D.H., Gibson K.M., Le T.P., Ribes A., Rodriguez de Cordoba S.,
RA Ugarte M., Penalva M.A.,
RT "The molecular basis of 3-methylcrotonylglycinuria, a disorder of
RT leucine catabolism.";
RL Am. J. Hum. Genet. 68:334-346(2001).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT MCGI PHE-454.
RX MEDLINE=21295033; PubMed=11401427;
RA Obata K., Fukuda T., Morishita R., Abe S., Asakawa S., Yamaguchi S.,
RA Yoshino M., Ihara K., Murayama K., Shigenoto K., Shimizu N., Kondo I.,
RT "Human biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase
RT gene (MCCA): cDNA sequence, genomic organization, localization to
RT chromosomal band 3q27, and expression.";
RL Genomics 72:145-152(2001).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT MCGI PHE-535.
RX MEDLINE=21399419; PubMed=11406611;
RA Holzinger A., Roeschinger W., Lagier F., Mayerhofer P.U., Lichtner P.,
RA Kattenfeld T., Thuy L.P., Nyhan W.L., Koch H.G., Muntau A.C.,
RA Roscher A.A.;
RT "Cloning of the human MCCA and MCCB genes and mutations therein reveal
RT the molecular cause of 3-methylcrotonyl-CoA: carboxylase
RT deficiency";
RL Hum. Mol. Genet. 10:1299-1306(2001).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANTS MCGI VAL-289; SER-385; PRO-437 AND
RP HIS-532.
RX PubMed=11181649;
RA Baumgartner M.R., Almasanu S., Suormala T., Obie C., Cole R.N.,
RA Packman S., Baumgartner E.R., Valle D.;
RT "The molecular basis of human 3-methylcrotonyl-CoA carboxylase
RT deficiency";
RL J. Clin. Invest. 107:495-504(2001).
RN [5]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;
RT "NEDD human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
RN [6]
RP SEQUENCE FROM N.A.
TISSUE=Skeletal muscle;
MEDLINE=22388257; PubMed=12477932;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3) (-) = ADP +
CC phosphate + 3-methylglutaconyl-CoA.

```

```

CC      -!- COPACFACTOR: Biotin.
CC
CC      -!- PATHWAY: Leucine catabolism.
CC
CC      -!- SUBUNIT: Probably a dodecamer composed of six biotin-containing
CC      alpha subunits and six beta subunits.
CC
CC      -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC
CC      -!- DISEASE: Defects in MCCC1 are the cause of 3-
CC      methylcrotonylglycinuria type I (MCGI) [MIM:210200]; also
CC      designated CGA or CG2. MCGI is a recessive disease that is
CC      characterized by muscular hypotonia and atrophy, probably of
CC      spinal origin.
CC
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.eib-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC

```

[illegible]

Search completed: March 24, 2004, 22:43:03
Job time : 20 secs

FT	1	47	MITOCHONDRION (POTENTIAL).
TRANSIT	48	725	METHYLCROTONYL-COA CARBOXYLASE ALPHA CHAIN.
FT			
NP_BIND	209	214	ATP (POTENTIAL).
ACT_SITE	339	339	BY SIMILARITY.
BINDING	681	681	BIOTIN (BY SIMILARITY).
DOWAIN	538	541	POLY-SER.
DOWAIN	713	718	POLY-GLU.
VARIANT	289	289	A -> V (in MCGI; mild form).
FT			/FtId=VAR_012785.
FT	325	325	M -> R (in MCGI).
FT			/FtId=VAR_012786.
VARIANT	385	385	R -> S (in MCGI; severe form).
FT			/FtId=VAR_012787.
VARIANT	437	437	L -> P (in MCGI; severe form).
FT			/FtId=VAR_012788.
VARIANT	464	464	P -> H.
FT			/FtId=VAR_012789.
VARIANT	532	532	D -> H (in MCGI; severe form).
FT			/FtId=VAR_012790.
VARIANT	535	535	S -> F (in MCGI; asymptomatic form).
FT			/FtId=VAR_012791.
CONFLICT	469	469	F -> L (IN REF. 3 AND 4).
SEQUENCE	725 AA;	80433 MW;	B847C6B80E0606EC0 CRG64;

Query Match 15.7%; Score 911; DB 1; Length 725;